Ultraproducts and Łoś's Theorem: A Category-Theoretic Analysis

by

Mark Jonathan Chimes

Dissertation presented for the degree of Master of Science in Mathematics in the Faculty of Science at Stellenbosch University

Department of Mathematical Sciences,
University of Stellenbosch

Supervisor: Dr Gareth Boxall

March 2017

Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Signature: ....
Mark Jonathan Chimes

Date: March 2017

Abstract

Ultraproducts and Łoś's Theorem:
A Category-Theoretic Analysis

Mark Jonathan Chimes

Department of Mathematical Sciences,
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.

Dissertation: MSc

2017

Ultraproducts are an important construction in model theory, especially as applied to algebra. Given some family of structures of a certain type, an ultraproduct of this family is a single structure which, in some sense, captures the important aspects of the family, where "important" is defined relative to a set of sets called an ultrafilter, which encodes which subfamilies are considered "large". This follows from Łoś's Theorem, namely, the Fundamental Theorem of Ultraproducts, which states that every first-order sentence is true of the ultraproduct if, and only if, there is some "large" subfamily of the family such that it is true of every structure in this subfamily. In this dissertation, ultraproducts are examined both from the standard model-theoretic, as well as from the category-theoretic view. Some potential problems with the category-theoretic definition of ultraproducts are pointed out, and it is argued that these are not as great an issue as first perceived. A general version of Łoś's Theorem is shown to hold for category-theoretic ultraproducts in general. This makes use of the concept of injectivity of a (compact) tree, which is intended to generalize truth of first-order formulae (under given assignments of variables), and, in the category of relational structures, corresponds exactly to first-order formulae. This type of thinking leads to a means of characterizing fields in the category of rings, and a new proof that every ultraproduct of fields is a field, which takes place entirely in the category of rings (along with the inclusion of the category of fields). Finally, the family of all (category-theoretic) ultraproducts on a given family is shown to arise from the "codensity monad"

of the functor which includes the category of finite families into the category of families. In this sense, it is shown that ultraproducts are a rather natural construction category-theoretically speaking.

Ultraprodukte en Łoś se Stelling: 'n Category-Teoretiese Analiese

Mark Jonathan Chimes

Department Wiskunde Wetenskappe,
Universiteit Stellenbosch,
Privaatsak X1, Matieland 7602, Suid Afrika.

Thesis: MSc

2016

Ultraprodukte is 'n belangrike konstruksie in modelteorie, veral in hul toepassings in algebra. Gegewe 'n reeks strukture van 'n sekere tipe, is 'n ultraproduk van hierdie reeks 'n enkele struktuur wat, op 'n manier, die belangrikste aspekte van die reeks bevat, waar "belangrik" hier gedefinieër word met betrekking tot 'n versameling reekse wat 'n ultrafilter genoem word. Hierdie ultrafilter verteenwoordig watter subreekse deur die ultraproduk as "groot" beskou word. Dit is 'n gevolgtrekking van Łoś se Stelling, dit wil sê, 'n eerste-orde stelling is waar met betrekking tot die ultraproduk as, en slegs as, daar 'n "groot" subreeks (van die hoofreeks) bestaan sodat die stelling waar is met betrekking tot elke struktuur in dié subreeks. In hierdie tesis word ultraprodukte uit die standarde model-teoretiese oogpunt behandel, sowel as uit die oogpunt van kategorie teorie. Potentieële probleme met die kategorie-teoretiese ultraproduk word uitgelig, maar dit word geargumenteer dat hul nie so 'n groot probleem veroorsaak as wat dit blyk nie. 'n Algmene weergawe van Łoś se stelling is bewys vir alle kategorieë. Dít maak gebruik van die konsep van injektiwiteit van 'n (kompakte) boom. Die bedoeling hiervan is om die waarheid van 'n eerste-orde stelling (onder 'n gegewe toedeling van veranderlikes) te veralgemeen. Hierdie idee ly tot 'n metode om liggame in die kategorie van groepe uit te lig, sowel as 'n nuwe bewys dat elke ultraproduk van liggame weer self 'n liggaam is. Hierdie bewys neem heeltemaal in die kategorie van groepe plaas (tesame met die funktor wat die kategorie van liggame in die kategorie van groepe insluit). Laastens, word dit angevoer dat die reeks van alle (kategorie-teoretiese) ultraprodukte van 'n gegewe reeks bestaan uit die "codigtheids monade" van die funktor wat die kategorie van eindige reekse insluit in die kategorie van oneindige reekse. Hierdie is dan 'n oortuiging dat ultraprodukte redelik natuurlik bestaan, ten minste uit die oogpunt van kategorie-teorie.

Acknowledgements

I would firstly like to thank my supervisor, Dr Gareth Boxall, for putting up with me throughout this process. He has had to deal with my constant switching of topics and changing of approaches. He has listened to my sometimes incoherent rambling in the process of trying to understand the relevant mathematics. He has given up many hours meticulously combing through many error-ridden pages. Throughout all of this, he has managed to remain friendly and instructive. Gareth, you are an excellent supervisor, and have gone far beyond what you were strictly required to do.

Thank you, also, to Prof. Zurab Janelidze, who provided a great deal of advice and encouragement. Thank you for being patient, even when my ideas were a bit strange, or even downright wrong. Our many sessions discussing mathematics, whether it was category theory or simply philosophy, have been vital to my success.

Next, I would like to thank my girlfriend, Lilley. She has been there through all my moments of despondence when the thesis just seemed to loom over my head. She has celebrated with me whenever I completed each chapter, and comforted me when I faced the next. Lilley, you have really been very supportive.

To my parents, I extend a most heartfelt gratitude. Though they had to feign interest whenever I mentioned my work, they nevertheless did all they could to help me. I do not think that I would have stood a chance without their assistance both financially and emotionally. Mom, dad, thank you. I love you.

I spent countless hours on the phone with my brother, Liam, just chatting late into the night when I should have been sleeping or working on my thesis. Both he and my sister, Juanelle, were great at distracting me when I should be working. But more importantly, they were great for conversation and spending time when I wasn't working. My time at home helped refresh me when I was emotionally drained, and without them my life would have felt a bit empty.

To the Wolfpack, you know who you are! You guys have been great company over the past year, and are the people with whom I've spent most of my free time. Thanks for being patient with me, and for being my friends.

To the Guild. You guys are great! Sorry the thesis took my time away from you. I'll see more of you next year!

Thank you to the Frank-Wilhelm Bursary Fund for funding my masters degree. I am sincerely grateful for their help. Without them, I would not have been able to afford to study.

Finally, I would like to thank the University of Stellenbosch for providing the facilities and the opportunity for my studies, and my entire student experience. I will sincerely miss this university and its campus. Pectora roborant cultus recti.

Dedications

This thesis is dedicated to my girlfriend's guide-dog, who understands almost as much about my thesis as she does, but is a much better listener. She has guided me through the hard times I have experienced in my writing; Teska-Monster, you truly are a human's best friend.

Contents

Chapter 1

Preliminaries

1.1 Category Theory

1.1.1 Concepts Assumed and Described

This dissertation assumes a basic understanding of category theory including at least the following concepts:

- Definition of a category
- Basic examples of categories (Sets, Groups, Rings, Fields etc.)
- Functors and natural transformations
- Products of categories
- Comma categories
- Limits and Colimits
- Adjunctions

The following concepts are defined and described in this dissertation.

- Ends (§1.1.5)
- Monads (§1.1.4.1)
- Codensity Monads (§8)
- Category of Families (§8.10)
- Trees as per [AN79] (§6.2).

1.1.1.1 Further Reading

The classic text for Category Theory, is [ML98]. Additionally, [Bor94] (along with parts 2 and 3) is very good for reference. An older text cited in this dissertation is [HS73], though some of the terminology used is somewhat out of date. The reference [JHS05] is more modern and is available freely online.

1.1.2 Notation

- Categories are denoted in blackboard bold font, e.g. $\mathbb{C}, \mathbb{D}$.
- Objects and functors are generally denoted with uppercase letters (in math font), e.g. A, F.
- Morphisms
  - Morphisms are generally denoted with lowercase letters (in math font), e.g., e, f, or with lowercase greek letters, e.g. κ, λ.
  - The notation $f : A \to B$ denotes that f is a morphism with domain A and codomain B.
  - If $f : A \to B$ and $g : B \to C$, then the **composite** of f and g is denoted gf or sometimes $g \circ f$.
- Multiple objects from the same category, or multiple morphisms, functors, natural transformations etc. with the same domain and codomain, may be represented at once with a separation by commas. E.g. $A, B, C \in \mathbb{C}$ denotes that $A \in \mathbb{C}$ and $B \in \mathbb{C}$ and $C \in \mathbb{C}$. Similarly, $f, g : A \to B$ denotes $f : A \to B$, and $g : A \to B$, and similarly for $F, G : \mathbb{A} \to \mathbb{B}$.
- Functors
  - Notation for functors is similar to that for morphisms ($F : \mathbb{A} \to \mathbb{B}$, composition GF etc.)
  - Application of a functor F to an object A is denoted $F(A)$ or FA.
- Natural Transformations
  - Natural transformations are generally denoted with lowercase Greek letters, e.g. α, η.
  - The notation $\alpha : F \Rightarrow G$ denotes that α is a natural transformation from functor F to functor G.
  - Given functors $F, G : \mathbb{A} \to \mathbb{B}$ and natural transformation $\alpha : F \Rightarrow G$, the component of α at an object $A \in \mathbb{A}$ is denoted α_A, and is a morphism $\alpha_A : F(A) \to G(A)$.

- Limits
  - Given a functor $F : \mathbb{C} \to \mathbb{D}$, the limit of F is denoted $\varprojlim F$.
  - The limit above may also be denoted $\varprojlim_{C \in \mathbb{C}} F_C$, where $F_C = F(C)$. (This notation is somewhat ambiguous since it does not display the morphisms of the limit).
  - Projection maps out of a limit are generally denoted with a π.
  - A projection map from a product $\prod_{x \in X} A_x$ to the xth component A_x is denoted π_x.
  - The morphism from an object A induced into the product $\prod_{x \in X} A_x$ by the morphisms $f_x : A \to A_x$ is denoted $\langle f_x \rangle_{x \in X} : A \to \prod_{x \in X} A_x$.
  - Given $X \subseteq Y$, the morphism $\langle \pi_x \rangle_{x \in X} : \prod_{y \in Y} A_y \to \prod_{x \in X} A_x$ induced by $\pi_x : \prod_{y \in Y} A_y \to A_x$ for each $x \in X$ is denoted π^Y_X.
- Colimits
  - Given a functor $F : \mathbb{C} \to \mathbb{D}$, the colimit of F is denoted $\varinjlim F$.
  - This colimit may also be denoted $\varinjlim_{C \in \mathbb{C}} F_C$, where $F_C = F(C)$.
  - Coprojections into a colimit are denoted by $\amalg_x : A_x \to \varinjlim_{y \in X} A_y$. (This notation is not standard).
- The LaTeX code for $\amalg$ is

```
\newcommand\rotpi{\rotatebox[origin=c]{180}{$\pi$}}
\newcommand\copi{\reflectbox{\rotpi}}
```

  Then use `\copi` in the document. This solution is taken from

  `http://tex.stackexchange.com/questions/303040/`
  `is-there-an-upside-down-version-of-the-symbol-for-pi/303042`

- An adjunction with F left adjoint to G is denoted $F \dashv G$ or $G \vdash F$.

1.1.3 Subobjects

One category-theoretic definition of subobjects is as follows.

Definition 1.1.3.1.
Given an object C in a category $\mathbb{C}$, a **subobject** of C is a pair $\langle C', m \rangle$ consisting of an object $C' \in \mathbb{C}$ and a monomorphism $m : C' \hookrightarrow C$. $\triangleleft$

Loosely speaking, the object C' may be referred to as the subobject.

The more precise version of this definition requires a certain equivalence class of morphisms.

Definition 1.1.3.2. Given an object $C \in \mathbb{C}$ and the collection $\mathrm{Hom}(*, C)$ of all morphisms with codomain C, define an equivalence relation $\equiv$ on $\mathrm{Hom}(*, C)$ as follows. Two morphisms $f, g \in \mathrm{Hom}(*, C)$ are equivalent $f \equiv g$ if, and only if, f factors through g and g factors through f, i.e., if there is some h such that $fh = g$ and there is some i such that $gi = f$. $\triangleleft$

That this forms an equivalence relation is routine to check.

$$\begin{tikzcd}
& C & \\
C' \arrow[ur, "f"] \arrow[rr, "i"] & & C'' \arrow[ul, "g"'] \arrow[ll, bend left, "h"]
\end{tikzcd}$$

The following definition of subobjects of C may be found in,for example, [MR77].

Definition 1.1.3.3.
A **subobject of** C is an equivalence class of $\mathrm{Hom}(*, C)$. $\triangleleft$

The domain of a representative of such an equivalence class may itself be referred to as a subobject of C. Usually, when the structure is concrete, the specific representative is chosen which corresponds to its image in C. For example, if the structures are groups, then a subgroup C' of a group C is considered to consist of elements from the supergroup C.

Examples of subobjects are the expected ones: subsets of sets, subgroups of groups, subspaces of topological spaces, and so forth.

1.1.4 Monoids and Monads

The technical definition of a monad is a monoid object in the category of endofunctors on a category $\mathbb{X}$.

A monoid $\langle M, \cdot, e\rangle$ may be viewed as consisting of a set M, a morphism $\cdot : M \times M \to M$ and a morphism $e : \{*\} \to M$ from a singleton that "picks out the unit", satisfying associativity and identity conditions. It should be noted that a singleton acts as an identity for the product $\times$ in the sense that $\{*\} \times X \cong X$ for any set X.

This idea may be abstracted with a different object M, other "product" $\otimes$ and other unit $e : I \to M$ which no longer has domain a singleton, but something else that acts as an identity for $\otimes$. These should still satisfy some form of identity and associativity conditions. Such a structure is called a **monoid object**.

A **monad** is a monoid object where M is an endofunctor, $\otimes$ is functor composition, and I is an identity functor. Each monad also has some multiplication μ and unit η (which are natural transformations since those are the morphisms in a category of functors) satisfying associativity and identity conditions.

The following definitions mostly follow [ML98], and the reader is encouraged to read there if they would like more information.

1.1.4.1 Monoids

A monoid object is only defined in the context of a **monoidal category**. Such a category is equipped with a bifunctor which works like multiplication, and an object which works as a unit for that multiplication.

Because the definition of a monoidal category is somewhat non-obvious (in particular since it relies on the "pentagon" and "triangle" coherence diagrams), the definition of a monoid object will be given first.

The most natural example of a monoidal category is the category of sets equipped with its product bifunctor $\times : \mathbf{Set} \times \mathbf{Set} \to \mathbf{Set} : (X, X) \mapsto X \times X$ and the unit object any singleton $\{*\}$. In this monoidal category, monoid objects are, in fact, exactly monoids equipped with the usual monoidal multiplication. The reader is encouraged to keep this example in mind as they read the definition of monoid object.

Definition 1.1.4.1. Given a monoidal category $\langle \mathbb{X}, \otimes, I \rangle$, a **monoid** in $\mathbb{X}$ is a triple $\langle M, \eta, \mu \rangle$ consisting of

- an object $M \in \mathbb{X}$
- a morphism $\eta : I \to M$ called the **unit** of the monoid
- a morphism $\mu : M^2 \to M$ called the **multiplication** of the monoid

satisfying the following two laws:

- Associative Law: $\mu\,(1_M \otimes \mu) = \mu\,(\mu \otimes 1_M)$
- Unit Law: $\mu\,(\eta \otimes 1_M) = 1_M = \mu\,(1_M \otimes \eta)$

$$\begin{tikzcd} M^3 \arrow[r, "1_M \otimes \mu"] \arrow[d, "\mu \otimes 1_M"'] & M^2 \arrow[d, "\mu"] \\ M^2 \arrow[r, "\mu"] & M \end{tikzcd} \qquad \begin{tikzcd} I \otimes M \arrow[r, "\eta \otimes I"] \arrow[dr, "1_M"'] & M^2 \arrow[d, "\mu"] & M \otimes I \arrow[l, "1_M \otimes \eta"'] \arrow[dl, "1_M"] \\ & M & \end{tikzcd}$$

Here, e.g., $1_M \otimes \mu$ is the morphism from $M \otimes (M \otimes M)$ to $M \otimes M$ which is the image under the bifunctor $\otimes$ of the morphism

$$\langle 1_M, \mu\rangle : \langle M, M \otimes M\rangle \to \langle M, M\rangle$$

in the product category $\mathbb{X} \times \mathbb{X}$. $\lhd$

Remark. Projections do not exist in an arbitrary monoidal category. $\lhd$

Remark. In an arbitrary monoidal category, the objects $M \otimes (M \otimes M)$ and $(M \otimes M) \otimes M$ might not be equal. The definition of a monoidal category is such that, although they may not be equal, the above objects are **isomorphic**. Hence, for example, the object $M \otimes M \otimes M$ may be defined up to isomorphism.

Similarly, although $I \otimes M$, M, and $M \otimes I$ are not equal, they are isomorphic.

The specific isomorphisms making these objects isomorphic form part of the definition of a given monoidal category, though they were not displayed in the triple $\langle \mathbb{X}, \otimes, I\rangle$ above. It should technically be a tuple $\langle \mathbb{X}, \otimes, I, \alpha, \lambda, \rho\rangle$, where α, λ and ρ are natural isomorphisms defined below. $\lhd$

1.1.4.2 Monoidal Categories

Definition 1.1.4.2. A **monoidal category** is a tuple $\langle \mathbb{X}, \otimes, I, \alpha, \lambda, \rho\rangle$, (often written simply as a triple $\langle \mathbb{X}, \otimes, I\rangle$) consisting of

- n category $\mathbb{X}$
- an object $I \in \mathbb{X}$ called the **unit object**
- a bifunctor $\otimes : \mathbb{X} \times \mathbb{X} \to \mathbb{X}$ called the **monoidal product**
- Natural isomorphisms
  - α with component $\alpha_{A,B,C} : (A \otimes B) \otimes C \cong A \otimes (B \otimes C)$ for each $A, B, C \in \mathbb{X}$, called the **associator**.
  - λ with component $\lambda_X : I \otimes X \cong X$ called the **left unitor**.
  - ρ with component $\rho_X : X \otimes I \cong X$ called the **right unitor**.

$\lhd$

Furthermore, the above satisfies the following coherence conditions. The following two diagrams are commutative (note that the bifunctor $\otimes$ is written in infix notation):

The pentagon diagram:

$$\begin{tikzcd}[column sep=small]
& ((A\otimes B)\otimes C)\otimes D \arrow[dl, "(\alpha_{A,B,C})\otimes 1_D"'] \arrow[dr, "\alpha_{(A\otimes B),C,D}"] & \\
(A\otimes(B\otimes C))\otimes D \arrow[d, "\alpha_{A,(B\otimes C),D}"'] & & (A\otimes B)\otimes(C\otimes D) \arrow[d, "\alpha_{A,B,(C\otimes D)}"] \\
A\otimes((B\otimes C)\otimes D) \arrow[rr, "1_A\otimes(\alpha_{B,C,D})"'] & & A\otimes(B\otimes(C\otimes D))
\end{tikzcd}$$

and the triangle diagram:

$$\begin{tikzcd}
(A\otimes I)\otimes B \arrow[rr, "\alpha_{A,I,B}"] \arrow[dr, "\rho_A\otimes 1_B"'] & & A\otimes(I\otimes B) \arrow[dl, "1_A\otimes\lambda_B"] \\
& A\otimes B &
\end{tikzcd}$$

1.1.4.3 Monads

Definition 1.1.4.3. Given an endofunctor $F : \mathbb{X} \to \mathbb{X}$ on a category $\mathbb{X}$, the nth composite F^n is

$$F^n = \underbrace{F \cdots F}_{n \text{ times}}.$$

$\lhd$

Definition 1.1.4.4. Given categories $\mathbb{X}, \mathbb{Y}$, and $\mathbb{Z}$, and functors $F, G : \mathbb{X} \to \mathbb{Y}$ and $H : \mathbb{Y} \to \mathbb{Z}$, as well as a natural transformation $\alpha : F \Rightarrow G$, the natural transformation $H\alpha$ is defined at each $X \in \mathbb{X}$ as

$$(H\alpha)_X = H(\alpha_X).$$

$$\begin{tikzcd}
\mathbb{X} \arrow[r, bend left, "F"{name=U}] \arrow[r, bend right, "G"'{name=D}] \arrow[Rightarrow, from=U, to=D, "\alpha"] & \mathbb{Y} \arrow[r, "H"] & \mathbb{Z}
\end{tikzcd}$$

Similarly, given categories $\mathbb{W}, \mathbb{X}$, and $\mathbb{Y}$, and functors $E : \mathbb{W} \to \mathbb{X}$ and $F, G : \mathbb{X} \to \mathbb{Y}$, as well as a natural transformation $\alpha : F \Rightarrow G$, the natural transformation αE is defined at each $W \in \mathbb{W}$ as

$$(\alpha E)_W = (\alpha_{E(W)}).$$

$$\begin{tikzcd} \mathbb{W} \arrow[r, "E"] & \mathbb{X} \arrow[r, bend left, "F"{name=U}] \arrow[r, bend right, "G"'{name=D}] & \mathbb{Y} \arrow[from=U, to=D, Rightarrow, "\alpha"] \end{tikzcd}$$

◁

Definition 1.1.4.5. Given a category $\mathbb{X}$, a **monad** in $\mathbb{X}$ is a triple $\langle M, \eta, \mu \rangle$ consisting of

- an endofunctor $M : \mathbb{X} \to \mathbb{X}$
- a natural transformation $\eta : 1_{\mathbb{X}} \Rightarrow M$ called the **unit** of the monad
- a natural transformation $\mu : M^2 \Rightarrow M$ called the **multiplication** of the monad

satisfying the following two laws:

- Associative Law: $\mu(M\mu) = \mu(\mu M)$
- Unit Law: $\mu(\eta M) = 1_M = \mu(M\eta)$

$$\begin{tikzcd} M^3 \arrow[r, "M\mu"] \arrow[d, "\mu M"'] & M^2 \arrow[d, "\mu"] \\ M^2 \arrow[r, "\mu"] & M \end{tikzcd} \qquad \begin{tikzcd} 1_{\mathbb{X}}M \arrow[r, "\eta M"] \arrow[dr, "1_M"'] & M^2 \arrow[d, "\mu"] & M1_{\mathbb{X}} \arrow[l, "M\eta"'] \arrow[dl, "1_M"] \\ & M & \end{tikzcd}$$

◁

Proposition 1.1.4.6. A monad is exactly a monoid object in the monoidal category $\langle \mathbf{End}(\mathbb{X}), 1_{\mathbb{X}}, \circ \rangle$ of endofunctors on a given category $\mathbb{X}$.

Here, the unit object is given by the identity functor $1_{\mathbb{X}}$, and the monoidal product is the bifunctor $\circ$ which takes a pair of endofunctors $\langle F, G \rangle$ to their composite $G \circ F$.

Proof. The corresponding natural isomorphisms α, λ and ρ are trivial (since composition of functors is associative, and composition with identity leaves a functor unchanged). This is the statement that the endofunctors on $\mathbb{X}$ form a **strict monoidal category**.

The proof that these endofunctors actuallys form a monoidal category is then a routine check of the definitions, as is the proof that monads correspond exactly to monoid objects in this monoidal category. □

Remark. In older literature, monads are referred to as **triples**. ◁

Monads arise from adjunctions. In fact, for any monad, it is possible to find an adjunction from which that monad can be derived.

Theorem 1.1.4.7. Given an adjunction $\langle F, G, \eta, \epsilon\rangle$, (where η is the unit and ϵ the co-unit of the adjunction, respectively)

$$\begin{tikzcd} \mathbb{X} \arrow[rr, bend left, "F"] \arrow[rr, phantom, "\perp"] & & \mathbb{A} \arrow[ll, bend left, "G"] \end{tikzcd},$$

the triple $\langle GF, \eta, G\epsilon F\rangle$ is a monad, where $G\epsilon F = (G\epsilon)F = G(\epsilon F)$.

Proof. See [ML98, Chapter 6] for the details. □

Theorem 1.1.4.8. Given a monad $\langle M, \eta_M, \mu\rangle$, there is an adjunction $\langle F, G, \eta, \epsilon\rangle$ such that $M = GF$, $\eta_M = \eta$, and $\mu = G\epsilon F$.

Proof. See [ML98, Chapter 6]. □

1.1.4.4 Example of Monad

Define

$$M : \mathbf{Set} \to \mathbf{Set} : X \mapsto \mathcal{P}(X)$$

with unit $\eta : 1_{\mathbf{Set}} \Rightarrow M$

$$\eta_X : X \to \mathcal{P}(X) : x \mapsto \{x\}$$

and multiplication $\mu : M^2 \to M$

$$\eta_X : \mathcal{P}(\mathcal{P}(X)) \to \mathcal{P}(X) : S \mapsto \cup S$$

where $\cup S$ is the union of sets in S:

$$\cup S = \cup_{Y \in S} Y = \{y | y \in Y, Y \in S\}.$$

It is routine to check that this is a monad.

1.1.5 Ends

1.1.5.1 Wedge

Definition 1.1.5.1. Given categories $\mathbb{C}$ and $\mathbb{X}$, a bifunctor

$$S : \mathbb{C}^{\mathrm{op}} \times \mathbb{C} \to \mathbb{X},$$

and an object $W \in \mathbb{X}$, a **wedge** from W to S is a family

$$\langle \omega_B : W \to S(B,B) \rangle_{\langle B \in \mathbb{C} \rangle}$$

of morphisms such that for any pair of objects $B, C \in \mathbb{C}$ and any morphism $f : B \to C$, it holds that

$$S(1_B, f)\, \omega_B \quad = \quad S(f, 1_C)\, \omega_C$$

as in the following diagram:

$$\begin{tikzcd}
 & S(B,B) \arrow[dr, "S(1_B,f)"] & \\
W \arrow[ur, "\omega_B"] \arrow[dr, "\omega_C"'] & & S(B,C) \\
 & S(C,C) \arrow[ur, "S(f,1_C)"'] &
\end{tikzcd}$$

◁

1.1.5.2 Definition of End

Definition 1.1.5.2. Given categories $\mathbb{C}$ and $\mathbb{X}$, and a bifunctor

$$S : \mathbb{C}^{\mathrm{op}} \times \mathbb{C} \to \mathbb{X},$$

the **end** of S, denoted

$$\int_{B \in \mathbb{C}} S(B,B)$$

is a pair $\langle X, \chi \rangle$ consisting of an object $X \in \mathbb{X}$ and a wedge χ from X to S which is universal, in the sense that for any other pair consisting of an object W and family of morphisms

$$\langle \omega_B : W \to S(B,B) \rangle_{B \in \mathbb{B}}$$

with the relevant commutative properties, there is a unique morphism

$$h : W \to X$$

such that, for every $B \in \mathbb{B}$, it holds that

$$\chi_B \circ h = \omega_B,$$

as in the following diagram:

$$\begin{tikzcd}
 & & S(B,B) \arrow[dr, "S(1_B, f)"] & \\
W \arrow[r, dashed, "h"] \arrow[urr, bend left, "\omega_B"] \arrow[drr, bend right, "\omega_C"'] & X \arrow[ur, "\chi_B"] \arrow[dr, "\chi_C"'] & & S(B,C) \\
 & & S(C,C) \arrow[ur, "S(f, 1_B)"'] &
\end{tikzcd}$$

$\triangleleft$

Remark. As is usual for universal constructions, the end of S is unique up to isomorphism. Additionally, similarly to limits, the symbol $\int_{B \in \mathbb{C}} S(B, B)$ represents the object of the end, and, by abuse of language, is itself referred to as the end. $\triangleleft$

Definition 1.1.5.3. Given categories $\mathbb{C}$ and $\mathbb{X}$, and a bifunctor

$$S : \mathbb{C}^{\mathrm{op}} \times \mathbb{C} \to \mathbb{X},$$

the **coend** of S, denoted

$$\int^{B \in \mathbb{C}} S(B, B)$$

is defined dually to the end. $\triangleleft$

1.1.5.3 Example of End

Given a category $\mathbb{B}$, and functors $F, G : \mathbb{B} \to \mathbf{Set}$, into the category of sets, the bifunctor

$$\prod_{F(-)} G(-) : \mathbb{B}^{\mathrm{op}} \times \mathbb{B} \to \mathbf{Set}$$

maps an object (B, C) to the product $\prod_{F(B)} G(C)$.

In §8.3.2, it is shown that (the object for) the end of this functor is the set of natural transformations from F to G

$$\int_{B \in \mathbb{B}} \prod_{F(B)} G(B) \quad = \quad \mathrm{Nat}(F, G).$$

1.2 Set Theory

1.2.1 Overview

The level of set theory used in this dissertation is rather light, and is of a standard level for many mathematical texts.

In particular, a metamathematical universe of sets is assumed to exist which satisfies ZFC, the Zermelo-Fraenkel Axioms with the Axiom of Choice. All categories, models, mathematical structures, etc. used in this dissertation are assumed to be specific examples of such sets.

Remark. In particular, the above implies, for example, that the category of all sets is itself a set. This paradox is dealt with in §1.2.3 below. $\triangleleft$

The reader is welcome to read, for example, [Sho77], or [BM77][Chapter 10] for an example of the ZFC axiomatization and a further description of set theory.

1.2.2 Notation

Sets may be specified using the **set-builder** notation, e.g., $Y = \{x \in X \mid \phi(x)\}$ specifies the set Y is defined as the subset of X such that the (first-order) property ϕ is true for each x. The colon version $Y = \{x \in X : \phi(x)\}$ may also be used. The set $\mathbb{N}$ denotes the natural numbers $\{1, 2, \ldots\}$ and is usually assumed to start from 1, but usage may be inconsistent. For a particular application, it may be emphasised that $\mathbb{N}$ starts from 1 by using the notation $\mathbb{N}^*$, or that it starts from 0 by using the notation $\mathbb{N}_0$.

A function f from X to Y is written $f : X \to Y$. If f takes each element x to an element y_x depending on x. This may be written

$$f : X \to Y : x \mapsto y.$$

Function application is written with parentheses $f(x) = y_x$, including for n-ary functions $f(x_1, \ldots, x_n) = y_{x_1,\ldots x_n}$.

The set Y^X is the set of all functions from X to Y.

Tuples are written with angular brackets $\langle x_1, \ldots, x_n \rangle$. Similarly, families are written $\langle x_i \rangle_{i \in I}$ where I is some indexing set, which will often be the natural numbers $\mathbb{N}$.

A family $\langle x_i \rangle_{i \in I} \in X^I$ is a function $f : I \to X : i \mapsto x_i$.

Families indexed by $\mathbb{N}$ may be written in tuple notation, e.g., as $\langle x_1, x_2, x_3, \ldots \rangle$. An ordered tuple $\langle x_1, \ldots, x_n \rangle$ can also be written as a family $\langle x_i \rangle_{i \leq n}$.

An n-ary relation R is a set of ordered tuples $\langle x_1, \ldots, x_n \rangle$, all of which have the same arity n.

Given, for example, a binary relation R, writing $R(x_1, x_2)$ is equivalent to writing $\langle x_1, x_2 \rangle \in R$.

Functions are a special case of relations, satisfying certain uniqueness and existence axioms. An n-ary function is a type of $(n+1)$-ary relation.

1.2.3 Cardinality

Questions of size are mostly ignored. It is assumed that everything used exists within some sufficiently large set-theoretic universe. So, for example, the category of sets is, in fact, a **large** category of **small** sets. The category of categories is the **even larger** category of **large** categories, or something similar.

In those cases where such an expansion to a large set may not be possible, the collection may be explicitly referred to as a **class**.

Additionally, questions involving the continuum hypothesis or anything similar will not feature, and the reader is welcome to make their own assumptions in this regard.

1.2.4 Axiom of Choice

The existence of the non-principal ultrafilters (defined in §3.3.3) necessary for non-trivial ultraproducts is dependent on the axiom of choice.

The use and requirements of this axiom with respect to ultrafilters is described in §3.3.3.

Since it so crucial for ultraproducts in any case, it seems inappropriate to be careful with its use in the rest of this dissertation. Hence, if the axiom is needed for other mathematical or meta-mathematical theorems, it will be used without necessarily alerting the reader to this fact.

For more information on the axiom of choice and its implications, the interested reader is advised to consult [Jec77], [Her06], or [BM77, Chapter 10].

1.3 Logic

1.3.1 Overview

It is assumed that the reader has at least a basic knowledge of mathematical logic. However, for completeness, definitions of languages, theories, models and so forth are given below.

Apart from this, the level of model-theory used in the paper is quite light, and that which is necessary (namely, ultraproducts) is defined in detail.

For further reference, the reader is encouraged to consult the various sources available in the compilation [Bar77a] which is a quite exhaustive text on logic. In particular, in this compilation, [Bar77b] provides a basic introduction to logic and [Ekl77] a basic introduction model theory.

Alternatively, for a detailed text in logic (including model theory), the reader may consult [BM77]. For a greater focus on (introductory) model theory, there is [Mar02], though it contains few references to ultraproducts. For a more ultraproduct-focussed approach, the reader may consult [BS71].

Finally, at the intersection of category theory and logic, there is [MR77].

1.3.2 Logic Preliminaries

Definition 1.3.2.1. A **string** is a word, (a finite tuple), and a **character** is a singleton string. Two strings written next to each other represents concatenation. $\triangleleft$

Definition 1.3.2.2. The set of **truth values** is a set

$$\mathbb{T} = \{\mathbf{True}, \mathbf{False}\}$$

where **True** should be interpreted as truth and **False** should be interpreted as falsehood or contradiction.

By definition, $\neg\mathbf{True} = \mathbf{False}$ and $\neg\mathbf{False} = \mathbf{True}$. $\triangleleft$

1.3.3 Languages

Definition 1.3.3.1. A **language** or **signature** $\mathcal{L}$ is a triple

$$\langle \langle \Sigma_n^R \rangle_{n \in \mathbb{N}_0}, \langle \Sigma_n^f \rangle_{n \in \mathbb{N}_0}, \Sigma^c \rangle$$

consisting of the relation, function and constant symbols respectively.

The relation symbols consists of a family $\langle \Sigma_n^R \rangle_{n \in \mathbb{N}_0}$, where each Σ_n^R is a set of symbols, representing the "n-ary" relation symbols. Similarly for the function symbols. All constants are nullary, so Σ^c is the set of constant symbols. $\triangleleft$

Definition 1.3.3.2. We may write a (countable) signature as

$$\langle \rho_1, \rho_2, \ldots, \zeta_1, \zeta_2, \ldots, \kappa_1, \kappa_2, \ldots \rangle$$

provided that the type and arity of each symbol is made clear. $\triangleleft$

1.3.4 Structures

Definition 1.3.4.1. Fix a language $\mathcal{L}$

A **model-theoretic structure in** $\mathcal{L}$ (or simply a **structure**) is a tuple

$$\mathcal{A} = \langle A, R^{\mathcal{A}}, f^{\mathcal{A}}, c^{\mathcal{A}} \rangle$$

where A is a set and $R^{\mathcal{A}}$ is a family

$$R^{\mathcal{A}} = \left\langle R_n^{\mathcal{A}} \right\rangle_{n \in \mathbb{N}_0}$$

of functions

$$R_n^{\mathcal{A}} : \Sigma_n^R \to \mathcal{P}\left(A^n\right)$$

which associates with each n-ary relation symbol ρ a corresponding n-ary relation

$$R_n^{\mathcal{A}}(\rho) \subseteq A^n.$$

similarly, $f^{\mathcal{A}}$ is a family of functions

$$f_n^{\mathcal{A}} : \Sigma_n^f \to A^{(A^n)}$$

associating to each n-ary function symbol ζ a function $f_n^{\mathcal{A}}(\zeta) : A^n \to A$, and $c^{\mathcal{A}}$ is a function that associates to each constant symbol $\kappa \in \Sigma^c$ an element $c^{\mathcal{A}}(\kappa) \in A$. $\triangleleft$

Definition 1.3.4.2. Given two structures $\langle A, R^{\mathcal{A}}, f^{\mathcal{A}}, c^{\mathcal{A}} \rangle$ and $\langle B, R^{\mathcal{B}}, f^{\mathcal{B}}, c^{\mathcal{B}} \rangle$, a **morphism of structures**, **homomorphism of structures**, or simply **homomorphism**

$$m : \langle A, R^{\mathcal{A}}, f^{\mathcal{A}}, c^{\mathcal{A}} \rangle \to \langle B, R^{\mathcal{B}}, f^{\mathcal{B}}, c^{\mathcal{B}} \rangle$$

is a map $m : A \to B$ which preserves the relations, functions and constants in the sense that

- For any n-ary relation symbol $\rho \in \Sigma_n^R$, for $a_1, \ldots, a_n \in A$,
$$\langle a_1, \ldots, a_n \rangle \in R_n^{\mathcal{A}}(\rho) \quad \Rightarrow \quad \langle m(a_1), \ldots, m(a_n) \rangle \in R_n^{\mathcal{B}}(\rho).$$
- For any n-ary function symbol $\zeta \in \Sigma_n^f$,
$$m(f_n^{\mathcal{A}}(\zeta)(a_1, \ldots, a_n)) = f_n^{\mathcal{B}}(\zeta)(m(a_1), \ldots, m(a_n)).$$
- For any constant symbol $\kappa \in \Sigma^c$,
$$m(f^{\mathcal{A}}(\kappa)) = f^{\mathcal{B}}(\kappa).$$

$\triangleleft$

Remark. Each language gives rise to a different category (of the structures in that language). $\triangleleft$

1.3.5 Relational Structures

Constants and functions can be encoded via relation symbols as well. Any n-ary function is given by an $n+1$-ary relation (with certain properties) and any constant is nullary function given by a certain unary relation.

Remark. Notice that $n = 0$ is permitted. Hence this construction allows for languages with nullary relation symbols corresponding to atomic sentences. The interpretation $R_0^{\mathcal{A}}(\rho) \subseteq A^0$ is then either empty or the one element subset of A^0, which correspond to falsehood and truth respectively. $\lhd$

Definition 1.3.5.1. A **relational structure** is a model-theoretic structure

$$\mathcal{A} = \langle A, R^{\mathcal{A}}, f^{\mathcal{A}}, c^{\mathcal{A}} \rangle$$

such that the functions $f^{\mathcal{A}}$ and constants $c^{\mathcal{A}}$ are both empty. It may be written simply as $\mathcal{A} = \langle A, R^{\mathcal{A}} \rangle$. $\lhd$

Remark. Note that two relational structures may both be empty, and yet non-isomorphic, because $0^0 = 1$ and the interpretation of ρ may be false in the one but true in the other.

The inclusion of any constant symbol would force the structures to be non-empty, however. $\lhd$

Remark. The initial object in any category of relational structures is then the empty structure **with all relational symbols interpreted as empty**. $\lhd$

1.3.6 Propositional Logic

Definition 1.3.6.1. The **language of propositional logic** is a tuple

$$\langle \neg, \wedge, (,), \Omega \rangle,$$

where the first four components are symbols and the last is a set of symbols, and each component is interpreted as follows

- $\neg$ is "negation", i.e., "not".
- $\wedge$ is "conjunction", i.e., "and".
- $($, $)$ are parentheses for grouping.
- Ω is a set of atomic propositions.

$\lhd$

Definition 1.3.6.2. The set Φ^* **well-formed formulae of propositional logic** is defined inductively on the length of formulae.

- Each atomic proposition $P \in \Omega$ is a formula $P \in \Phi^*$.
- If $\psi \in \Phi^*$ then $\neg(\psi) \in \Phi^*$.
- If $\psi, \tau \in \Phi^*$ then $(\psi) \wedge (\tau) \in \Phi^*$.

$\triangleleft$

Remark. Here e.g. ψ represents a string and $\neg(\psi)$ represents the characters '$\neg$' , '(' followed by every character of the string ψ and finally the character ')'. $\triangleleft$

Definition 1.3.6.3. The symbols $\vee$ (disjunction "or") and $\rightarrow$ (implication) are defined as shorthand:

- $(\psi) \vee (\tau) = \neg((\neg(\psi)) \wedge (\neg(\tau)))$.
- $(\psi) \rightarrow (\tau) = (\neg(\psi)) \vee (\tau)$.

for any formulae ψ and τ. $\triangleleft$

Definition 1.3.6.4. If $\phi = P$ for some atomic proposition $P \in \Omega$, then ϕ is an **atomic formula**. $\triangleleft$

Definition 1.3.6.5. For well-formed formulae $\psi, \phi \in \Phi^*$, define ψ **to be a subformula of** ϕ, written $\psi \leq \phi$ if ψ is one of the formulae from which ϕ is inductively defined. $\triangleleft$

Definition 1.3.6.6. For an atomic proposition $P \in \Omega$ and formula ϕ, define **P occurs in** ϕ, written $P \in \phi$, if P is one of the formulae from which ϕ is inductively defined (i.e., if P is a subformula $P \leq \phi$). $\triangleleft$

Definition 1.3.6.7. Let Ω be a family of atomic propositional symbols.

An **atomic valuation** $\bar{v}$ is a map

$$\bar{v} : \Omega \rightarrow \mathbb{T}$$

taking each P to a corresponding truth value $t_P \in \mathbb{T}$. $\triangleleft$

Definition 1.3.6.8. Let Φ^* be the set of all well-formed propositional formulae, with atomic propositions in Ω.

Then a **valuation** v is a function

$$v : \Phi^* \to \mathbb{T}$$

taking each atomic proposition $P \in \Omega$ to a corresponding truth-value t_P, and defined recursively on (length of) formulae as follows. For $\psi, \tau \in \Phi^*$ define

- $v(\neg\psi) = \textbf{True}$ iff $v(\psi) = \textbf{False}$.
- $v(\psi \wedge \tau) = \textbf{True}$ if and only if both $v(\psi) = \textbf{True}$ and $v(\tau) = \textbf{True}$.

$\lhd$

Definition 1.3.6.9. If $\bar{v} : \Omega \to \mathbb{T}$ is an atomic valuation, and $v : \Phi^* \to \mathbb{T}$ is a valuation, then v **extends** $\bar{v}$ if, and only, if $v(P) = \bar{v}(P)$ for each atomic proposition $P \in \Omega$. $\lhd$

Definition 1.3.6.10. A family Σ of propositional formulae is **consistent** if there is some valuation

$$v : \Phi^* \to \mathbb{T}$$

such that every $\phi \in \Sigma$ is evaluated as true $v(\phi) = \textbf{True}$. $\lhd$

Definition 1.3.6.11. Fix a set V of characters called **variables**. $\lhd$

Definition 1.3.6.12. Given a language $\mathcal{L}$, define the terms of $\mathcal{L}$, written $\textbf{Terms}_\mathcal{L}$, recursively as follows:

- $x \in \textbf{Terms}_\mathcal{L}$ for every variable $x \in V$.
- $\kappa \in \textbf{Terms}_\mathcal{L}$ for every $\kappa \in \Sigma^c$
- if $t_1, \ldots, t_n \in \textbf{Terms}_\mathcal{L}$, then $\zeta(t_1, \ldots, t_n) \in \textbf{Terms}_\mathcal{L}$ for every $\zeta \in \Sigma^f_n$.

$\lhd$

Definition 1.3.6.13. Given a term $t \in \textbf{Terms}_\mathcal{L}$, the **variables of** t is defined as the set of variables of x which occur in t, i.e., those variables which occur at some stage in which t is built. $\lhd$

Definition 1.3.6.14. Define a set Ω of atomic propositions as follows:

- $(t_1 = t_2) \in \Omega$ for each pair of terms $t_1, t_2 \in \textbf{Terms}_\mathcal{L}$.

- $\rho(t_1, \ldots, t_n) \in \Omega$ for each n-ary relation symbol ρ.

$\triangleleft$

From hereon, assume that the set of atomic propositions Ω is equal to the above.

Definition 1.3.6.15. Given a proposition $P \in \Omega$, the set of **variables of** P is defined as those variables which occur in P, i.e., the variables of each term from which P is built.

If a variable x occurs in P, it may be written $x \in P$. $\triangleleft$

Definition 1.3.6.16. Given a propositional formula $\phi \in \Phi^*$, the set of **variables occuring in** ϕ is defined as

$$\{x \in V \mid x \in P, \text{ for some atomic proposition } P \in \phi\}$$

The number of variables occurring in ϕ is the cardinality of the above set. $\triangleleft$

1.3.7 First-Order Logic

Definition 1.3.7.1. The **language of predicate logic**, or of **first-order logic** is a tuple

$$\langle \neg, \wedge, (,), \exists, V \rangle,$$

where the first four components are symbols and the last is a set of symbols, and where each component is interpreted as follows

- $\neg$, $\wedge$,(, and) are interpreted as in the language of propositional logic (Definition 1.3.6.1).
- $\exists$ is existence.
- V is a set of variables.

$\triangleleft$

Remark. The terms "first-order" and "predicate" will be used interchangeably in this dissertation. $\triangleleft$

Definition 1.3.7.2. The set Φ **well-formed formulae of predicate logic** is defined inductively on the length of formulae.

- Each atomic proposition $P \in \Omega$ is a formula $P \in \Phi$.

- If $\psi \in \mathbf{\Phi}$ then $\neg(\psi) \in \mathbf{\Phi}$.
- If x occurs in ψ and $\exists x(\tau)$ does not occur in ψ for any formula τ, then $\exists x(\psi) \in \mathbf{\Phi}$.
- If $\psi, \tau \in \mathbf{\Phi}^*$, and if, for each variable $x \in \psi$, there is no formula ϵ such that $\exists x(\epsilon)$ occurs in τ, and similarly, if for each $x \in \tau$, there is no ϵ such that $\exists x(\epsilon)$ occurs in ψ then $(\psi) \wedge (\tau) \in \mathbf{\Phi}$.

$\lhd$

Remark. The specification for conjunction is somewhat complicated in order to avoid the problem of variable capture.

Usually in a text on logic, a formula such as $(\exists x P(x)) \wedge Q(x)$, where P and Q are atomic propositions, is allowed. The two occurrences of the variable x are treated differently. The first is considered "captured" by the existence quantifier, whereas the second is free.

In this text, a different approach is taken. A formula such as the above is simply disallowed. One must use two different variables, say x_1 and x_2. Hence the restriction on forming conjunctions.

Essentially, all of the formulae that could usually be formed are still possible; if one wished to form a conjunction of two formulae in which some variables are problematic, it is possible to simply replace one of them with an equivalent formula using different variables in the quantifiers.

Now, this restriction is relaxed to allow formulae such as $(\exists x P(x)) \wedge Q(x)$, but this is understood as a notational convenience, rather than strictly accurate.

Hence, for example,

$$(\exists x(R_1(x))) \wedge R_2(x)$$

may be written using the same variable x in both cases, with the understanding that it could be rewritten unambiguously as

$$(\exists x(R_1(x_1))) \wedge R_2(x_2)$$

if necessary. $\lhd$

Definition 1.3.7.3. The symbols $\vee, \rightarrow, \forall, \exists_{x_1,\ldots,x_n}$ and so forth are defined as shorthand:

- $\vee$ and $\rightarrow$ are defined as for predicate logic. (Definition 1.3.6.3)
- $\forall x(\tau)$ is shorthand for $\neg(\exists x(\neg(\tau))$ for any formula τ.

- $\exists_x(\tau)$ is equivalent to $\exists x(\tau)$ for any formula τ, and similarly for $\forall$.
- $\exists_{x_1,\ldots,x_n}(\tau)$ is shorthand for $\exists_{x_1}(\exists_{x_2}(\cdots(\exists_{x_n}(\tau))\cdots))$, for any formula τ, and similarly for $\forall$.

$\triangleleft$

Remark. The symbols $\forall$ and $\exists$ are called quantifiers. $\triangleleft$

Definition 1.3.7.4. In this section, care is taken with parentheses. However, in general they may be omitted where it will be unambiguous to do so.

In order to ensure unambiguity, operations will bind via the following priority:

1. Parentheses
2. Function Application
3. Relation and Equality
4. Universal Quantification
5. Existential Quantification
6. Conjunction
7. Disjunction
8. Implication
9. Left-to-right

$\triangleleft$

So, for example, the following expressions are equivalent:

$$\forall_x\ x \cdot y = z \qquad \equiv \qquad \forall_x((x \cdot y) = z),$$

and so are the following following expressions:

$$\begin{aligned} &\exists_x\ R(x) \wedge R(y) \vee R(z) \to R(c) \\ \equiv \quad &\Bigg(\bigg(\Big(\big(\exists_x(R(x))\big) \wedge \big(R(y)\big)\Big) \vee \big(R(z)\big)\bigg) \to \big(R(c)\big)\Bigg). \end{aligned}$$

Also, when unambiguous, square brackets may be used in place of parentheses, so e.g. $\forall_x[R(x)]$ means the same as $\forall_x(R(x))$.

Since conjunction is associative when considered alone, and the same holds for disjunction, a chain of such expressions may be written without ambiguity, e.g.

$$P_1 \wedge P_2 \wedge P_3 \wedge \cdots \wedge P_n.$$

In addition, since equality is transitive, multiple chained equalities may be written without ambiguity, e.g.

$$x_1 = x_2 = \cdots = x_n,$$

which, formally, means

$$(x_1 = x_2) \wedge (x_2 = x_3) \wedge \cdots \wedge (x_{n-1} = x_n).$$

Definition 1.3.7.5. Any formula ϕ written in the language of first-order logic, but which contains no quantifiers (i.e., the formula $\exists(x)(\tau)$ is not a subformula of ϕ for any variable x and formula τ), is a **quantifier free formula** and is also referred to as a **propositional formula**. $\triangleleft$

Definition 1.3.7.6. A variable $x \in \phi$ is a **captured variable of** ϕ if there is some formula τ such that $\exists_x(\tau)$ is a subformula of ϕ.

The **free variables** of a formula ϕ are those variables $x \in \phi$ which are not captured variables. $\triangleleft$

Definition 1.3.7.7. Given a first-order formula $\phi \in \boldsymbol{\Phi}$, the set of **free variables occuring in** ϕ is defined as

$$\text{free}(\phi) = \{x \in \phi \mid x \text{ is free in } \phi\}.$$

The number of free variables in ϕ, written $|\phi|$, is the cardinality of the above set. $\triangleleft$

Definition 1.3.7.8. A formula ϕ is a **sentence** if it contains no free variables. $\triangleleft$

Definition 1.3.7.9. A propositional formula ϕ may be written

$$\phi = \phi(x_1, \ldots, x_n)$$

to indicate that the variables $x_1, \ldots, x_n$ are a superset of the set of free variables of ϕ. (Note this means that the variables may be displayed even if they are not free in ϕ or indeed even if they do not occur in ϕ.) $\triangleleft$

1.3.8 Models and Theories

Definition 1.3.8.1. Given a language $\mathcal{L}$ with variables V, a structure $\mathcal{A}$ of $\mathcal{L}$, and a subset $V' \subseteq V$, an **assignment of variables** $\bar{a}$ is a function $\bar{a} : V' \to A$. $\lhd$

Definition 1.3.8.2. Given an assignment of variables $\bar{a}$, the corresponding **assignment of terms** a is the function a : **Terms** $\to A$ defined recursively on (length of) terms as follows. For terms $t_1, \ldots, t_n, t \in$ **Terms** with $|t_i| < t$ for each $i \leq n$, the assignment $a(t)$ is defined

- $a(x) = \bar{a}(x)$ for variables $x \in V$.
- $a(\kappa) = a(c^{\mathcal{A}}(\kappa))$
- $a(\zeta(t_1, \ldots, t_n)) = f_n^{\mathcal{A}}(\zeta)(a(t_1), \ldots, a(t_n))$.

$\lhd$

Definition 1.3.8.3. Given an assignment $\bar{a}$ of variables, and the set **Terms** of terms of $\mathcal{L}$, the **term valuation** $\bar{v}_{\bar{a}}$ under $\bar{a}$ is an atomic valuation (Definition 1.3.6.7) defined as follows.

- $\bar{v}_{\bar{a}}(t_1 = t_2) =$ **True** iff $a(t_1) = a(t_2)$, for terms $t_1, t_2 \in$ **Terms**.
- $\bar{v}_{\bar{a}}(\rho(t_1, \ldots, t_n)) =$ **True** iff $\langle a(t_1), \ldots, a(t_n) \rangle \in R_n^{\mathcal{A}}(\rho)$, for terms $t_1, \ldots, t_n \in$ **Terms**.

$\lhd$

Definition 1.3.8.4. Given a language $\mathcal{L}$, and an assignment of variables $\bar{a} : V \to \mathbb{T}$, the **valuation** $v_{\bar{a}}$ of the set $\mathbf{\Phi}_n$ of all **first-order** formulae in n free variables under the assignment a is defined recursively on (length of) first-order formulae as follows. For $\phi \in \mathbf{\Phi}_n$:

- If ϕ is atomic, then $v_{\bar{a}}(\phi) = \bar{v}_{\bar{a}}(\phi)$ is the term valuation.
- $v_{\bar{a}}(\neg\psi) =$ **True** iff $v_{\bar{a}}(\psi) =$ **False**.
- $v_{\bar{a}}(\psi \wedge \tau) =$ **True** iff $v_{\bar{a}}(\psi) =$ **True** and $v_{\bar{a}}(\tau) =$ **True**.
- $v_{\bar{a}}(\exists_x \psi(x)) =$ **True** if, and only if, there is some assignment $\bar{a}'$ such that $\bar{a}'(y) = \bar{a}(y)$ for all $y \neq x$, and such that, for the valuation $v_{\bar{a}'} : \mathbf{\Phi}_n \to \mathbb{T}$ under the assignment $\bar{a}'$ it holds that $v_{\bar{a}'}(\psi(x)) =$ **True**.

◁

Definition 1.3.8.5. Given a formula ϕ whose free variables are a subset of some set $\mathrm{free}(\phi) \subseteq V' \subseteq V$, the valuation $v_{\bar{a}}$ under some assignment $\bar{a}$ is defined as above.

If $v_{\bar{a}}(\phi) = \mathbf{True}$, then the structure $\mathcal{A}$ is said to **model ϕ under the assignment** (of variables) $\bar{a}$, written

$$\mathcal{A} \models \phi[\bar{a}].$$

This is also written as

$$\mathcal{A} \models \phi(x_1, \ldots, x_n)[\langle a_1, \ldots, a_n \rangle],$$

where $\bar{a}(x_i) = a_i$ for each $i \leq n$.

Parentheses also may be used instead of square brackets, with the meaning obvious from whether the symbols between the parentheses are variables (such as x_i) or specific elements (such as a_i). ◁

In order to prevent captured variables causing any interference, we will assume any set Σ of first-order formulae satisfies the following property:

Definition 1.3.8.6. A set Σ of first order formulae is **captured non-clashing** if, for any $\psi, \tau \in \Sigma$, no free variable of ψ is a captured variable of τ. ◁

Remark. This above definition was created for this dissertation and is not in general use. ◁

Definition 1.3.8.7. Given a (captured non-clashing) set Σ of first order formulae ϕ, define the **free variables of Σ** to be

$$\bigcup_{\phi \in \Sigma} \{x \mid x \text{ free in } \phi\}$$

◁

Definition 1.3.8.8. Given a set Σ of first order formulae ϕ, a set of variables V' containing the free variables of Σ, and a structure $\mathcal{A}$, as well as an assignment of variables $\bar{a} : V' \to \mathcal{A}$, then $\mathcal{A}$ models Σ under the assignment $\bar{a}$, written

$$\mathcal{A} \models \Sigma[\bar{a}],$$

if, for each $\phi \in \Sigma$, it holds that

$$\mathcal{A} \models \phi[\bar{a}].$$

◁

Definition 1.3.8.9. A set Σ of first-order formulae is **satisfiable** if there is some model $\mathcal{A}$ and some assignment of variables $\bar{a} : V' \to \mathcal{A}$ such that

$$\mathcal{A} \models \Sigma[\bar{a}].$$

$\lhd$

Remark. If a set Σ of first-order formulae consists only of sentences, then the free variables of Σ are empty, and hence the empty assignment $\emptyset : \emptyset \to \mathcal{A}$ is a valid assignment for any $\mathcal{A}$.

Note also that for any $\mathcal{A}$, the truth of $\mathcal{A} \models \Sigma[\emptyset]$ will be the same as the truth of $\mathcal{A} \models \Sigma[\bar{a}]$ for any assignment $\bar{a}$ of variables. $\lhd$

Definition 1.3.8.10. If the empty assignment is being used, we may write

$$\mathcal{A} \models \Sigma \qquad \text{in place of} \qquad \mathcal{A} \models \Sigma[\emptyset].$$

$\lhd$

Definition 1.3.8.11. If a set Σ of sentences is satisfiable, it is also called **consistent**. $\lhd$

Remark. The definition of consistency may also be given by saying that there is no proof of both ϕ and $\neg\phi$ from Σ for any sentence ϕ.

By Theorem 1.3.11.1, both definitions are equivalent. $\lhd$

Definition 1.3.8.12. A **Theory $\mathcal{T}$ of $\mathcal{L}$** is a consistent set of first-order sentences written using the symbols in the language $\mathcal{L}$. $\lhd$

Remark. Some authors assume $\mathcal{T}$ is **deductively closed**, i.e., that if some statement ϕ is provable from $\mathcal{T}$, then it is an element of $\mathcal{T}$. This is not assumed in this dissertation. $\lhd$

Definition 1.3.8.13. Given a language $\mathcal{L}$ and a Theory $\mathcal{T}$ of $\mathcal{L}$, a structure $\mathcal{A}$ in $\mathcal{L}$ is said to **model $\mathcal{T}$**, or to **be a model of $\mathcal{T}$**, written

$$\mathcal{A} \models \mathcal{T}$$

if every sentence of $\mathcal{T}$ is true when interpreted in $\mathcal{A}$. $\lhd$

1.3.9 Theory of Groups

As an example, the language of groups may be written $\mathcal{L} = \langle\ \cdot\ ,\ e\ ,\ ^{-1}\ \rangle$, where $\cdot$ is a binary function (written in infix notation) representing multiplication, e is a constant representing the identity, and $^{-1}$ is a unary function (written in postfix notation) representing inverse.

The theory of groups may then be written

- Associativity of Multiplication: $\forall_{x,y,z}([(x \cdot y) \cdot z] = [x \cdot (y \cdot z)])$.
- Behaviour of Identity: $\forall_x(e \cdot x = x = x \cdot e)$.
- Behaviour of Inverse: $\forall_x(x \cdot x^{-1} = e = x^{-1} \cdot x)$.

This is far from the only means of describing the theory of groups.

1.3.10 Groups as Relational Structures

For example, the language of groups can be captured by the unary relation $\mathrm{Id}(x)$ for the "x is identity", the binary relation $\mathrm{Inv}(x, y)$ for "x is left-inverse to y" and the ternary relation $*(x, y, z)$ for "$xy = z$".

The category of groups is then equivalent to the full subcategory of this category consisting of those objects that satisfy the following axioms:

- Functionality of star: $\forall_{x,y}\exists!_z(*(x, y, z))$
- Behaviour of Identity: $\forall_x[\mathrm{Id}(x) \to \forall_y(*(x, y, y) \wedge *(y, x, y))]$
- Existence of Identity: $(\exists_x \mathrm{Id}(x))$
- Behaviour of Inverse: $\forall_{x,y}[\mathrm{Inv}(x, y) \leftrightarrow \forall_z(\mathrm{Id}(z) \to *(x, y, z))]$
- Existence of Inverse: $\forall_x\exists_y \mathrm{Inv}(x, y) \wedge \mathrm{Inv}(y, x)$
- Associativity: $\forall_{x,y,z,u,v}[*(x, y, u) \wedge *(y, z, v) \to \exists_w(*(u, z, w) \wedge *(x, v, w))]$

Remark. That the inverse and identity are in fact functions follows from the fact that they are unique, which itself follows from the above group axioms.
$\lhd$

Remark. Strictly speaking, only one-sided identities and inverses need to be defined, since the existence of the other-sided identities and inverses then follows, as well as the fact-that each left identity/inverse is equal to the right.
$\lhd$

That this forms a full subcategory means that the morphisms between the relational structures that are groups correspond exactly to group homomorphisms.

The same can be done to describe the category of any model-theoretic structure (Definition 1.3.4.1) as a subcategory of the category of relational structures with the corresponding signature. Such structures include rings and fields, but not topological spaces.

The above formalization does not account for multi-sorted structures, but it should not be difficult to extend it to do so.

1.3.11 Gödel's Theorems and Other Important Theorems of Logic

There are important theorems of completeness and incompleteness due to Gödel. We will not go into detail here. However, the reader is encouraged to read [Bar77b], [BM77, Chapters 1 and 2], or a similar source for more information.

The completeness theorem for first-order logic is based on making use of some system of deduction rules and tautologies for first-order logic. Such a system is not described here, but examples of such systems may be found in both [Bar77b] and [BM77].

Given such a system $\mathcal{S}$, the soundness and completeness theorems for first-order logic are as follows:

Theorem 1.3.11.1 (Completeness of First-Order Logic)**.** For any (well-orderable) language $\mathcal{L}$, for any first-order theory $\mathcal{T}$ in $\mathcal{L}$, and any sentence ϕ in $\mathcal{L}$, the following two statements are equivalent (model existence):

- There is no proof of $\neg\phi$ from $\mathcal{T}$ using $\mathcal{S}$,
- $\mathcal{A} \models \phi$ for some model $\mathcal{A} \models \mathcal{T}$.

and the following two statements are equivalent (completeness):

- There is a proof of ϕ from $\mathcal{T}$ using $\mathcal{S}$.
- $\mathcal{A} \models \phi$ for every model $\mathcal{A} \models \mathcal{T}$.

$\lhd$

An important, related theorem is the compactness theorem.

Theorem 1.3.11.2. Given a set Σ of first-order formulae, the following two statements are equivalent:

- Σ is satisfiable (Definition 1.3.8.9)
- Σ is finitely satisfiable. I.e., every finite subset $\Sigma' \subseteq \Sigma$ is satisfiable.

$\lhd$

Both of the above theorems require the Axiom of Choice, or a weaker axiom which is sufficiently strong, such as the Boolean Prime Ideal Theorem (See §3.3.3 for details).

In fact, the above two theorems are inter-provable:

Theorem 1.3.11.3. Assume the Zermelo-Fraenkel axiom system ZF.

Then the following are equivalent.

- Theorem 1.3.11.1
- Theorem 1.3.11.2

$\lhd$

There is also the incompleteness theorem, which uses "complete" in a different sense to the above.

In particular, the theorem above uses the notion of **semantic completeness**, which is to say, every tautology ϕ (a sentence provable using $\mathcal{S}$ and $\mathcal{T}$) is true in every model.

The theorem below uses the notion of **syntactical completeness** which means that, for every sentence ϕ, either ϕ is provable from $\mathcal{T}$, or $\neg\phi$ is provable from $\mathcal{T}$. This is also called **deductive completeness** or **negation completeness**.

Let the logical system $\mathcal{S}$ be fixed.

Theorem 1.3.11.4 (Incompleteness of Peano Arithmetic)**.** There is no recursively enumerable theory $\mathcal{T}$ which encodes / proves Peano Arithmetic, which is consistent, and such that for each statement ϕ, either ϕ or $\neg\phi$ is provable from $\mathcal{T}$. $\lhd$

1.4 Well-Founded Induction

Well-Founded Induction, or **Nötherian Induction** as it is also known, is induction without a base-case, or, more accurately, induction with an implicit base-case.

A relation R is well-founded if we cannot "chain backwards" infinitely far.

Definition 1.4.0.1. Assume we have some set X on which a binary relation R is defined.

Then, the binary relation R is **well-founded** if, for any $x \in X$, there do not exist infinitely many $y_i \in X$ such that $\cdots Ry_{i_3}Ry_{i_2}Ry_{i_1}Rx$. $\lhd$

One example of a well-founded relation is the element-of relation $\in$ for sets.

Proposition 1.4.0.2 ((Binary) Well-founded Induction)**.** Let R be a well-founded, binary relation on a set X, and let $\phi(x)$ be a first-order formula in one free variable.

Then, if for every $x \in X$ the implication below holds

$$\left[\forall_{y\in X}(yRx \to \phi(y))\right] \to \phi(x)$$

then, for every $x \in X$, the statement $\phi(x)$ holds. ◁

As an example, let X be some 'small' set model of sets as per §1.2.3 (i.e., X is a set in the larger universe whose elements are themselves interpreted as all sets under consideration). let $R =\in$, and assume the hypothesis of the above proposition holds for $\in$. Then, in particular, $\phi(\emptyset)$ also holds: $\phi(x)$ is vacuously true for every element $x \in \emptyset$, and so, by assumption, it is true of $\emptyset$.

Then, by the hypothesis, it is true of $\{\emptyset\}$. And thus is also true of $\{\{\emptyset\}\}$ and $\{\emptyset, \{\emptyset\}\}$ and so on for every finite set. Then it is also true for infinite sets like $\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \ldots\}$. Hence, intuitively, it should hold for **every set**. This is exactly what the principle of well-founded induction on sets says.

A similar argument follows for any well-founded relation. By well-foundedness, every "backwards chain" eventually ends in some element x which is minimal with respect to the relation. In the case of sets, the minimal element is $\emptyset$ and is unique, but in general there may be arbitrarily many minimal elements. Nevertheless, the same type of reasoning leads one to conclude the truth of the well-founded induction principle.

The principle may be extended as follows.

Definition 1.4.0.3. For $n \in \mathbb{N}^*$, define $yR^n x$ to mean

$$\exists y_1, \ldots y_n x = y_1 R y_2 R \cdots R y_n R y.$$

◁

Proposition 1.4.0.4 (Generalized (Binary) Well-founded Induction)**.** Let R be a well-founded, binary relation on a set X, and let $\phi(x)$ be a first-order formula in one free variable.

Let $n \in \mathbb{N}^*$.

Then, if for every $x \in X$ the implication below holds

$$\left[\forall_{y\in X}(yR^n x \to \phi(y))\right] \to \phi(x)$$

then, for every $x \in X$, the statement $\phi(x)$ holds. ◁

Remark. In this case, the "base-cases" for the induction are all the $x \in X$ such that there is no $y \in Y$ with $yR^n x$. Such an x is not necessarily 'minimal', since even an x for which there is a y with $yR^{n-1}x$ still satisfies this criterion. $\triangleleft$

1.5 Varieties of Algebras

Definition 1.5.0.1. Fix a signature $\mathcal{L}$ consisting of only relation symbols. Let $\mathcal{A}$ be a system of axioms defined using $\mathcal{L}$. Then $\mathcal{A}$ is a a **system of universally quantified equations** if each axiom of $\mathcal{A}$ is a sentence of the form (or is equivalent to a sentence of the form)

$$\forall_{x_1,\ldots x_n}\Big[t_2(x_1,\ldots,x_n) = t_2(x_1,\ldots,x_n)\Big]$$

for some $n \in \mathbb{N}$ and terms $t_1, t_2 \in \mathbf{Terms}_{\mathcal{L}}$. $\triangleleft$

Definition 1.5.0.2. Let $\mathcal{L}$ be a language containing only relation symbols, and let $\mathcal{L}'$ be a language containing relation, function and constant symbols.

Let $\mathcal{A}$ be a system of axioms defined over $\mathcal{L}$ and let $\mathcal{A}'$ be a system of axioms defined over $\mathcal{L}'$.

Then $\mathcal{A}$ and $\mathcal{A}'$ are **equivalent** if

- There is a bijection $\beta : \mathcal{L}' \to \mathcal{L}$ mapping
  - each variable x to a variable y
  - each n-ary relation symbol R' to an n-ary relation symbol R.
  - each n-ary function symbol f to an $(n+1)$-ary relation symbol R_f.
  - each constant symbol c to a unary relation symbol R_c.
- It is a theorem of $\mathcal{A}$ that each image R_f of a function symbol f is the relation of a function (i.e., that $\forall_{x_1,\ldots,x_n}\exists!_{x_{n+1}}R_f(x_1,\ldots,x_{n+1})$).
- It is a theorem of $\mathcal{A}$ that each image R_c of a constant symbol c is the relation of a constant (i.e., that $\exists!_x R_c(x)$).
- Each axiom of $\mathcal{A}'$ interpreted in $\mathcal{L}$ is a theorem of $\mathcal{A}$, where interpretation is understood as replacing phrases in the following manner.

  Define $\gamma :$ inductively on terms and formulae as follows:
  - $\gamma(\phi \wedge \psi) = \gamma(\phi) \wedge \gamma(\psi)$
  - $\gamma(\neg(\phi)) = \neg\gamma(\phi)$
  - $\gamma(\exists_x \phi(x)) = \exists_{\beta(x)}\gamma(\phi(x))$

- $\gamma(x_1 = x_2) = (\beta(x_1) = \beta(x_2))$
- $\gamma(x = c) = \exists_{y_1}(R_c(y_1) \wedge (y_1 = \beta(x)))$
- $\gamma(t_1 = t_2) = \exists_{y_1,y_2}(\gamma(y_1 = t_1) \wedge \gamma(y_2 = t_2))$
- $\gamma(R'(t_1, \ldots, t_n)) = \exists_{y_1,\ldots,y_n}(R(y_1, \ldots, y_n) \wedge \gamma(y_1 = t_1) \wedge \cdots \wedge \gamma(y_n = t_n))$
- $\gamma(x = f(t_1, \ldots, t_n)) = \exists_{y_1,\ldots,y_n}(R_f(y_1, \ldots, y_n, \beta(x)) \wedge \gamma(y_1 = t_1) \wedge \cdots \wedge \gamma(y_n = t_n))$

(Where in each case the variables in the existence quantifier are chosen so as to not coincide with any other variables.) Each axiom ϕ of $\mathcal{A}'$ is then replaced with $\gamma(\phi)$.

- Each axiom of $\mathcal{A}$ interpreted in $\mathcal{L}'$ is a theorem of $\mathcal{A}'$. Again, there is a means of translating formulae, which is essentially the reverse of the above, but the details are not given here.

$\triangleleft$

Definition 1.5.0.3. Let $\mathcal{A}$ be a system of axioms, and let $\mathbb{C}$ be the class of (model-theoretic) structures satisfying $\mathcal{A}$. If there is some universally quantified $\mathcal{A}'$ equivalent to $\mathcal{A}$, then $\mathbb{C}$ is a **variety of algebras** (or **variety of universal algebras**). The category $\mathbb{C}$ of such structures is also called a **variety of algebras**. $\triangleleft$

Remark. A **variety of algebras** should not be confused with an **algebraic variety** from commutative algebra. $\triangleleft$

Remark. Some authors restrict varieties to contain only non-empty structures. In this dissertation, empty structures are permitted, but the theorems will remain correct regardless of the assumption. $\triangleleft$

The following trivial lemma is useful.

Lemma 1.5.0.4. The category $\mathbb{R}$ of relational structures (as defined in §1.3.5) over any language $\mathcal{L}$ is a variety of algebras.

Proof. An empty system of axioms is trivially a system of universally quantified equations. $\square$

By a theorem of Birkhoff, (see e.g., [BS81], available online) this is equivalent to the following condition

Theorem 1.5.0.5 (Birkhoff)**.** Fix a language $\mathcal{L}$. Let $\mathbb{M}$ be the class of models of $\mathcal{L}$ over the empty system of axioms. Let $\mathcal{A}$ be a system of axioms in $\mathcal{L}$ and let $\mathbb{C}$ be the **subclass** of $\mathbb{M}$ defined by these axioms.

Then the following are equivalent

1. $\mathbb{C}$ is a variety of universal algebras.
2. $\mathbb{C}$ is closed under **subobjects**, **products**, and **homomorphic images** taken in $\mathbb{M}$.

$\triangleleft$

The latter property is made more precise below. For the definition of subobjects, see §1.1.3.

Definition 1.5.0.6. Fix a language $\mathcal{L}$. Let $\mathbb{M}$ be the class of models of $\mathcal{L}$ over the empty system of axioms. Let $\mathcal{A}$ be a system of axioms in $\mathcal{L}$ and let $\mathbb{C}$ be the **subclass** of $\mathbb{M}$ defined by these axioms.

Then $\mathbb{C}$ is

- **closed under homomorphic images** (which can also be stated as **closed under surjective homomorphisms** or **closed under quotients**) if, given an object $C \in \mathbb{C}$, an object $C' \in \mathbb{M}$ and a surjective homomorphism $h : C \to C'$ of $\mathbb{M}$ it holds that
  - C' is an element of $\mathbb{C}$.
  - the map $h : C \to C'$ is a homomorphism when interpreted in $\mathbb{C}$.
- **closed under subobjects** if, given an object $C \in \mathbb{C}$, and a subobject $C' \subseteq C$ in $\mathbb{M}$, then
  - C' is an element of $\mathbb{C}$.
  - C' is a subobject of C in $\mathbb{C}$.
- **closed under products** if, for every product $\prod_{i \in I} C_i$ (including empty I) of objects $C_i \in \mathbb{C}$ calculated in $\mathbb{M}$,
  - that product is an element of $\mathbb{C}$.
  - that product is the product of $\langle C_i \rangle_{i \in I}$ calculated in $\mathbb{C}$.

$\triangleleft$

Chapter 2

Overview

2.1 Summary

2.1.1 Ultraproducts

Ultraproducts are a useful construction in model theory. In some sense, they allow one to capture the information of a family $\langle A_i \rangle_{i \in I}$ of mathematical (in particular, algebraic) structures A_i indexed by some set I, and condense that information into a single structure of the same form. For example, if $\langle A_i \rangle_{i \in \mathbb{N}}$ is a family of groups, then an ultraproduct of this family is itself a group, and any first-order sentence which is true of 'most' of the groups (in particular, 'most' will chosen so as to include any cofinite subfamily) is true of the ultraproduct.

2.1.2 Ultrafilters

In order to define precisely what 'most' means in terms of the ultraproduct, one needs the concept of an **ultrafilter** defined on the indexing set I. An ultrafilter $\mathcal{U}$ is a subset of the power-set $\mathcal{P}(I)$, and intuitively it represents the 'important' 'large' subsets of I. In particular, it chooses these sets in such a way that any first-order sentence ϕ is decidable in the ultraproduct. In particular, for any such ϕ, there is either some 'large set' $H \in \mathcal{U}$ for which ϕ holds for all of the 'large subfamily' $\langle A_i \rangle_{i \in H}$ indexed by H, or there is a large family in which $\neg\phi$ holds, and never both.

Furthermore, these ϕ are consistent, in that ψ and τ each hold in a large family if, and only if, $\psi \wedge \tau$ holds in some large family (and similarly for the other operations of first-order logic).

2.1.3 Łoś's Theorem

Note that the above is not the way that truth of sentences is **defined** in the ultraproduct, instead, the ultraproduct is defined by giving an interpretation of the symbols based on their interpretation in large subfamilies of the structures making up the ultraproduct.

Instead, it follows from the **Fundamental Theorem of Ultraproducts**, namely that the sentences ϕ which are true for the ultraproduct $\prod_{\mathcal{U}} A_i$ are those which are true on some large subfamily $\langle A_i \rangle_{i \in H}$. This fact is also called **Łoś's Theorem** after Jerzy Łoś.

ϕ true in the ultraproduct

$$\prod_{\mathcal{U}} A_i$$

if, and only, if

$$\cdots\cdots \underbrace{\cdots\cdots A_i \cdots\cdots}_{H} \cdots\cdots$$

ϕ true on a large subfamily

For any given set I, there are many possible ultrafilters $\mathcal{U} \subseteq \mathcal{P}(I)$. In particular there are many trivial **principal** ultrafilters which are of little use. However, the general existence of a non-principal ultrafilter, namely, one which contains all of the cofinite subsets of an infinite I (among other subsets), requires the axiom of choice. The use of such an ultrafilter ensures that at least the cofinite subfamilies are considered large. It then makes arbitrary, but consistent, choices for the rest. E.g., if the indexing set is $\mathbb{N}$, then a non-principal ultrafilter $\mathcal{U}$ contains any cofinite subset of $\mathbb{N}$, but also contains an arbitrary choice of whether the even numbers or the odd numbers are considered a large set. It is these many arbitrary, consistent choices which require the axiom of choice.

2.1.4 Model-Theoretic Ultraproduct

Model-theoretically, an ultraproduct of a family $\langle A_i \rangle_{i \in I}$ of model-theoretic structures (in some language $\mathcal{L}$, satisfying some theory $\mathcal{T}$) is defined as a certain product $\prod_{i \in I} A_i$ of the underlying sets of the family, quotiented out by a specific equivalence relation $\sim_{\mathcal{U}}$ dependent on the ultrafilter.

This equivalence relation equates two families $\langle a_i\rangle_{i\in I}, \langle b_i\rangle_{i\in I}, \in \prod_{i\in I} A_i$ if, and only if, they are equal on some large subset of I. The interpretation of $\mathcal{L}$ is also defined to be 'the same as' on large subsets. E.g., if $+ \in \mathcal{L}$, then the sum of two equivalence classes is defined

$$[\langle a_i\rangle_{i\in I}] + [\langle b_i\rangle_{i\in I}] = [\langle a_i + b_i\rangle_{i\in I}].$$

This is well-defined (independent of representatives), and furthermore, Łoś's Theorem implies that the ultraproduct so-defined will also satisfy whatever theory $\mathcal{T}$ is satisfied by the family, so that it is a structure of the same type.

2.1.5 Category-Theoretic Ultraproduct

The ultraproduct also has a category-theoretic definition. Given, a category $\mathbb{C}$ with all small products and (directed) colimits, a family $\langle A_i\rangle_{i\in I}$ of objects in $\mathbb{C}$, and an ultrafilter $\mathcal{U} \subseteq I$, the category-theoretic ultraproduct is the colimit of the products of large subfamilies (with morphisms the projection morphisms).

$$\begin{tikzcd}
 & W & & \\
 & & \prod\limits_{\mathcal{U}} A_x \arrow[ul, dotted, "h"'] & \\
\prod\limits_{x\in X} A_x \arrow[uur, bend left, "w_X"] \arrow[urr, "\rotatebox[origin=c]{180}{π}_X"'] \arrow[dr, "\pi^X_Z"'] & & & \prod\limits_{x\in Y} A_x \arrow[uull, bend right, "w_Y"'] \arrow[ul, "\rotatebox[origin=c]{180}{π}_Y"'] \arrow[dll, "\pi^Y_Z"] \\
 & \prod\limits_{x\in Z} A_x \arrow[uuu, bend left, "w_Z"] \arrow[uur, "\rotatebox[origin=c]{180}{π}_Z"'] & &
\end{tikzcd}$$

Remark. Being a colimit, the cone commutes with these morphisms, and for any other such object W and commuting cone, there is a unique morphism out of the ultraproduct making the two cones commute. $\triangleleft$

2.1.6 Problems With the Category-Theoretic Ultraproduct

In many categories of interest (namely varieties of universal algebra), this construction exists and is equal to the model-theoretic ultraproduct. However, it does not always exist. E.g., the necessary products do not exist for fields. Furthermore, even in a category where it does exist, it may not correspond with the model-theoretic ultraproduct.

This problem may be resolved by expanding to a much larger category, namely the category of models of the empty theory of $\mathcal{L}$ (this dissertation assumes all of the symbols are relational symbols. This category is then referred to as the category of relational structures). The category-theoretic ultraproduct can then be calculated there, and it will correspond to the model-theoretic version. Furthermore, Łoś's Theorem implies it will then also satisfy the same $\mathcal{T}$ as the objects of the original $\mathbb{C}$.

2.1.7 Łoś's Theorem in Any Category

In fact, a version of Łoś's Theorem is true in every category in which the category-theoretic ultraproducts can be defined. In those categories where the model-theoretic and category-theoretic ultraproducts coincide, this is simply the classical version of Łoś's Theorem. This version of Łoś's Theorem in the category of relational structures also allows the expansion and calculation described above. However, the interpretation of this theorem in other categories is a bit different.

2.1.8 Injectivity Trees to Represent First-Order Formulae

There is a means of translating first-order formulae into a category-theoretic structure called a **tree** and truth of the first-order formula in an object A into a category-theoretic property called **injectivity of the object with respect to the tree**. In a category of relational structures, and in any variety of universal algebras, this corresponds exactly to the classical interpretation and truth of the formula under a given assignment of variables. However, this definition of trees, injectivity, and Łoś's Theorem make sense in any category.

2.1.9 Codensity Monads

The set $U(I)$ of ultrafilters $\mathcal{U} \subseteq \mathcal{P}(I)$ arises as a specific monad. Namely, the codensity monad for the functor which includes the category of finite sets into the category of sets.

Remark. The above fact has a great deal to do with the fact that the set of ultrafilters $U(I)$ is the Stone-Čech compactification of I with the discrete topology. This is not explored further in the current dissertation. ◁

The family

$$\left\langle \prod_{\mathcal{U}} A_i \right\rangle_{\mathcal{U} \in U(I)}$$

of category-theoretic ultraproducts of a given family $\langle A_i \rangle_{i \in I}$ also arises as a certain monad. Namely, the **codensity monad** for the functor which includes the category of finite families (families indexed by a finite set) into the category of all families.

2.2 Example Applications of Ultraproducts

2.2.1 Proof of Completeness of First-Order Logic

Ultraproducts can be used to simplify Gödel's proof of the completeness theorem for predicate calculus. [BS71, Chapter 12-1] contains the details (for sentences not containing equality).

2.2.2 Non-standard Model of Peano Arithmetic

Let $L = \langle S, +, \cdot \rangle$ be the language of arithmetic (where S is the symbol for the successor function $S(x) = x + 1$) and let T be a (recursive) axiomatization of Peano Arithmetic. (See [BM77][Ch 7, §9] for an example of such an axiomatization).

The natural numbers $\mathbb{N}$ under the usual interpretation of the language L is called the 'standard model of arithmetic'.

Let $\mathcal{U} \subseteq \mathcal{P}(\omega)$ be a non-principal ultrafilter on $\mathcal{P}(\omega)$. (The symbol ω is used to differentiate the indexing set from the model $\mathbb{N}$).

Form the ultraproduct

$$\mathbb{N}^\omega/\mathcal{U} = \prod_{\mathcal{U}} \mathbb{N}$$

By Łoś's theorem, this ultraproduct is elementarily equivalent to $\mathbb{N}$, i.e., it satisfies exactly the same first-order formulae. However, it also has elements which satisfy certain non-first-order properties that are not satisfied by any element of $\mathbb{N}$.

Definition 2.2.2.1. $S^n(0)$ is $SS\cdots S(0)$ where S is repeated n times. $S^0(0) = 0$. ◁

Definition 2.2.2.2. A **non-standard element of arithmetic** is an element which is neither 0 nor the n^{th} successor of 0. Put another way, x is nonstandard if $x \neq S^n(0)$ for any n. ◁

Proposition 2.2.2.3. The ultraproduct $\prod_{\mathcal{U}} \mathbb{N}$ has non-standard elements.

Proof. In the ultraproduct, the interpretation of the symbol 0 is the equivalence class $[\langle 0, 0, \ldots \rangle]$. Its successors are $[\langle 1, 1, \ldots \rangle]$, $[\langle 2, 2, \ldots \rangle]$, $[\langle 3, 3, \ldots \rangle]$ etc.

Consider the tuple $\langle 0, 1, 2, 3, \ldots \rangle$. Then it differs from the tuple $\langle k, k, \ldots \rangle$ on a cofinite and hence large set, for any $k \in \mathbb{N}$. Thus it cannot lie in any equivalence class $[\langle k, k, \ldots \rangle]$, and thus is not the n-th successor of $[\langle 0, 0, \ldots \rangle]$ for any n. □

Any such ultraproduct of $\mathbb{N}$ is uncountable. However, it is possible to obtain a non-standard, countable elementary substructure. See [BS71, Chapter 12] for the details.

2.2.3 Non-standard Reals

Similarly to the above, a non-standard model of the real numbers as a field can be constructed using ultraproducts.

The language of ordered fields is $L = \langle +, \cdot, 0, 1, < \rangle$. The structure of the real numbers $\mathbb{R}$ under the usual interpretation is called the standard model of the reals.

Again, let $\mathcal{U} \subseteq \mathcal{P}(\omega)$ be a non-principal ultrafilter on $\mathcal{P}(\omega)$ and form the ultraproduct

$$\mathbb{R}^\omega / \mathcal{U} = \prod_{\mathcal{U}} \mathbb{R}$$

This structure satisfies all of the same first-order formulae as the reals, but contains elements larger than any real number, and positive elements which are smaller than any positive real number (i.e., elements whose absolute value is smaller than the absolute value of any real number). These are called non-standard real numbers.

Proof. Consider the equivalence classes $[\langle 1, 2, 3, \ldots \rangle]$ and $[\langle \frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \ldots \rangle]$. Then for any $r \in \mathbb{R}$, $[\langle \frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \ldots \rangle] < [\langle r, r, r, \ldots \rangle] < [\langle 1, 2, 3, \ldots \rangle]$. □

Remark. The ultraproduct contains elements larger than any real number, but does not contain a largest element since the first-order formula $\exists_x \forall_y y < x$ is false in $\mathbb{R}$ and so must be false in the non-standard reals as well. Similarly, there is no smallest element. ◁

Remark. The non-standard reals contain numbers 'infinitely close' to any real, given by $[\langle r + \frac{1}{1}, r + \frac{1}{2}, r + \frac{1}{3}, \ldots \rangle]$. ◁

The details are given in [CK73].

The interested reader is encouraged to further consult [KS04] in which it is shown that there exists a definable, countably saturated elementary extension of the reals.

2.2.4 Field of Characteristic 0

Given only fields of characteristic p for any prime p, it is possible to prove there exists a field of characteristic 0.

Order the prime numbers from smallest to largest as $p_1 = 2, p_2 = 3, p_3 = 5, p_4 = 7, \ldots$. Let $\mathbb{F}_{p_i}$ denote the finite field of order p_i. So the family is $\langle \mathbb{F}_{p_i} \rangle_{i \in \omega}$. Let $\mathcal{U} \subseteq \mathcal{P}(\omega)$ be a non-principal ultrafilter and form the ultraproduct

$$\mathbb{F}_{p_i}^{\omega} / \mathcal{U} = \prod_{\mathcal{U}} \mathbb{F}_{p_i}$$

For any natural number n, the statement that $\mathbb{F}_{p_i}$ has characteristic greater than n is true for cofinitely many values i. Hence the ultraproduct must have characteristic greater than n for any n. Somewhat unintuitively, this does not imply that it has characteristic infinity, but that it has characteristic 0.

Remark. In some sense, this is because characteristics should not be ordered by size but by 'divisibility', with 1 being the "smallest" since it divides into everything, and 0 being the largest because everything divides into it. $\lhd$

In first-order language, the first order statement for a field having characteristic at most n for $n \geq 1$ is

$$\forall_x [(x = 0) \vee (x + x = 0) \vee \cdots \vee (nx = 0)]$$

where nx means $x + \cdots + x$ added n times.

The above ultraproduct does not satisfy this first-order statement for any n. By convention this is defined as the field having characteristic 0.

Remark. This field is an example of a pseudofinite, and hence quasifinite field. In fact, every non-principal ultraproduct of finite fields is pseudofinite. $\lhd$

See [Ax68] for more details on ultraproducts of finite fields, and on pseudo-finite, and quasi-finite fields.

2.2.5 Pseudofinite Groups

Definition 2.2.5.1. A group is called a **pseudofinite** group if it is elementarily equivalent to an ultraproduct of finite groups. $\lhd$

Pseudofinite groups have applications as a bridge between the finite and infinite model theory of groups. See [Wil95] as just one example.

2.2.6 Saturated Models

Of great importance in model theory, and the primary application of ultraproducts, is the concept of a **saturated model**.

It is described briefly here. The interested reader may look at [Mar02] for a description of saturated models, and at [CK73] and [Ekl77] for both a description and the means by which they are constructed using ultraproducts.

The concept of a saturated model requires the definition of a **type**, which is simply a (satisfiable) set of formulae. Intuitively a type may be seen as a description of possible elements of a model, or as a generalization of a conjunction of formulae.

The following definitions follow [Mar02].

Definition 2.2.6.1. Given a language $\mathcal{L}$ and theory $\mathcal{T}$ of $\mathcal{L}$, a **complete** n**-type** p of a theory $\mathcal{T}$ is a satisfiable set of first-order forumulae with n free variables in $\mathcal{L}$ such that $\mathcal{T} \subseteq p$ and such that for each formula ϕ of arity at most n in $\mathcal{L}$, either $\phi \in p$ or $\neg\phi \in p$. $\lhd$

Definition 2.2.6.2. Let $\mathcal{L}$ be a language, $\mathcal{T}$ a theory in $\mathcal{L}$, and $\mathcal{M}$ a model of $\mathcal{T}$.

Let $A \subset M$ be a subset of the underlying set of $\mathcal{M}$, and let $\mathcal{L}_A$ be the language obtained by adding each element of A as a constant symbol to $\mathcal{L}$, and let $\mathcal{T}_A$ be the corresponding extended theory.

Let κ be a cardinal. Then $\mathcal{M}$ is κ**-saturated**, if for

- every $A \subset M$ with $|A| < \kappa$,
- every $n \in \mathbb{N}$,
- and every complete n-type p of the language $\mathcal{L}_A$ and theory $\mathcal{T}_A$

the type p is realized in $\mathcal{M}$, i.e., there is some element $\langle a_1, a_2, \ldots a_n\rangle \in M^n$ such that every formula $\phi \in p$ holds for this tuple. $\lhd$

For example, the set $\{x > 0, x > 1, x > 2, \ldots\}$ is a type in the language of ordered fields. The axiom of choice may be used to form a complete type p containing this set. Hence, a sufficiently saturated model in the language of ordered fields will realize p.

Hence, taking the first-order axioms for the reals, and forming a saturated model of these axioms provides a model for the non-standard reals.

It is possible to obtain κ-saturated models by taking ultraproducts over κ-**good** ultrafilters, as defined in [CK73]. It may also be found in [Ekl77]. Any non-principal ultrafilter (§3.3.3) is sufficient for ω-saturation (where ω is the ordinal representing the natural numbers).

Chapter 3

Ultrafilters

3.1 Overview

Non-principal Ultrafilters make precise the notion of "large subsets" of a given set. An ultrafilter is a special type of filter, and hence filters are defined first.

Fix a set X. A filter is a subset of the power-set $\mathcal{P}(X)$ which is closed under intersections and supersets. A proper filter is one which does not contain the empty-set. By the superset condition, this is equivalent to the filter not being equal to the whole power-set.

An ultrafilter $\mathcal{U}$ is a (proper) filter which also satisfies the condition that for each subset $Y \in \mathcal{P}(X)$, either Y or its complement $X - Y$ is contained in $\mathcal{U}$.

This definition still allows for a somewhat trivial kind of ultrafilter, namely a "principal ultrafilter", which, given a specific element $x \in X$, consists of exactly those subsets $Y \in \mathcal{P}(X)$ such that $x \in Y$. In order to avoid this case, when making use of an ultrafilter, it is generally assumed that it is non-principal. However, principal ultrafilters are the only ultrafilters for a finite set X. Futhermore, even for infinite X, the existence of non-principal ultrafilters requires the axiom of choice (or something slightly weaker).

Many theorems of ultrafilters hold for both principal and non-principal cases, so the assumption of an ultrafilter being non-principal will generally be made explicit.

An ultrafilters satisfy a useful condition called the "partition condition". This is an alternate characterization of ultrafilters in the following sense. A subset $\mathcal{U} \in \mathcal{P}(X)$ is an ultrafilter if, and only if, for any partition of X into exactly three disjoint parts, exactly one of those parts is contained in the ultrafilter. A similar condition exists for every fixed number of parts greater than or equal to three.

Lastly, there is a monad, the ultrafilter monad, whose functor sends a set X to the set $U(X)$ of ultrafilters on X. This is a rather special monad. The functor

U is terminal amongst endofunctors on **Set** which preserve finite coproducts, and from this it follows that the monad structure on U is unique.

For references on the basic definitions of filters and ultrafilters, the reader is advised to consult [DP02]. The partition condition is first stated in [GH70], and the details of the ultrafilter monad may be found in [Bör87].

3.2 Filter

3.2.1 Definitions

Definition 3.2.1.1. Given a set X, a **filter** $\mathcal{F}$ on $\mathcal{P}(X)$ is a set of subsets $\mathcal{F} \subseteq \mathcal{P}(X)$ satisfying the following conditions. For $A, B \in \mathcal{P}(X)$:

1. If $A, B \in \mathcal{F}$ then $A \cap B \in \mathcal{F}$
2. If $A \in \mathcal{F}$ and $B \supseteq A$ then $B \in \mathcal{F}$

$\lhd$

Remark. The filter $\mathcal{F}$ is a sublattice of $\mathcal{P}(X)$ ordered by subset inclusion. $\lhd$

Definition 3.2.1.2. If $\mathcal{F} \neq \mathcal{P}(X)$ then the filter is called a **proper** filter, otherwise it is called an **improper** filter. $\lhd$

Note that $\mathcal{F} = \mathcal{P}(X)$ iff $\emptyset \in \mathcal{F}$.

Only proper filters are of interest. Hence many authors require that $\emptyset \notin \mathcal{F}$ for $\mathcal{F}$ to be a filter. Unless mentioned otherwise, it may be assumed all filters in this text are proper.

Remark. Filters can be defined generally on partially ordered sets, however for our purposes we only need them to be defined on the boolean algebra $\mathcal{P}(X)$ of subsets of a given set. $\lhd$

Remark. There is some ambiguity in the mathematical literature: A filter $\mathcal{F}$ on a poset (partially ordered set) X is $\mathcal{F} \subseteq X$, but a filter $\mathcal{F}$ on (sometimes over) a set X is $\mathcal{F} \subseteq \mathcal{P}(X)$. For clarity, we shall always define $\mathcal{F}$ as a filter on $\mathcal{P}(X)$. $\lhd$

The following is an especially important example of a filter.

Proposition 3.2.1.3. Let X be an infinite set and let $\mathcal{F} \subseteq \mathcal{P}(X)$ consist of all the cofinite sets of X.

$$\mathcal{F} = \{Y \in \mathcal{P}(X) : X - Y \text{ finite}\}$$

Then $\mathcal{F}$ is a (proper) filter on $\mathcal{P}(X)$.

Proof. Let $Y, Z \in \mathcal{F}$. Then $(X - Y)$ and $(X - Z)$ are both finite. Hence $(X - Y) \cup (X - Z) = X - (Y \cap Z)$ is finite and so $Y \cap Z \in \mathcal{F}$.

Let $Y \in \mathcal{F}$ and $Z \supseteq Y$. Then $(X - Z) \subseteq (X - Y)$ and since $X - Y$ is finite, $X - Z$ must also be. Hence $Z \in \mathcal{F}$. □

Definition 3.2.1.4. The above filter is called the **cofinite filter** or the **Frechét filter** on $\mathcal{P}(X)$. ◁

3.3 Ultrafilters

3.3.1 Definition

Definition 3.3.1.1. Given a set X, an **ultrafilter** $\mathcal{U}$ on $\mathcal{P}(X)$ is a set of subsets $\mathcal{U} \subseteq \mathcal{P}(X)$ satisfying the following conditions. For $A, B \in \mathcal{P}(X)$:

1. If $A, B \in \mathcal{U}$ then $A \cap B \in \mathcal{U}$
2. If $A \in \mathcal{U}$ and $B \supseteq A$ then $B \in \mathcal{U}$
3. Either $A \in \mathcal{U}$ or $X \backslash A \in \mathcal{U}$ but not both.

◁

Remark. By the above properties, an ultrafilter cannot contain the empty set, and is hence always a proper filter. If $X = \emptyset$ then there are no ultrafilters on $\mathcal{P}(X)$. ◁

Theorem 3.3.1.2. Let $x \in X$ and let $\mathcal{U} \subseteq \mathcal{P}(X)$ consist of all the subsets of X containing x.

$$\mathcal{U} = \{A \in X : x \in A\}$$

Then $\mathcal{U}$ is an ultrafilter on $\mathcal{P}(X)$.

Proof. For $A, B \in \mathcal{P}(X)$:

1. If $x \in A$ and $x \in B$ then $x \in A \cap B$.
2. If $x \in A$ and $B \supseteq A$ then $x \in B$.
3. Either $x \in A$ or $x \in (X - A)$ but not both.

□

Definition 3.3.1.3. An ultrafilter of the above form is referred to as a **principal ultrafilter** (the principal ultrafilter generated by x). ◁

3.3.2 Ultrafilters and Maximal Filters

The following propositions show that the ultrafilters are exactly the maximal filters.

Proposition 3.3.2.1. Let $\mathcal{U}$ be an ultrafilter on $\mathcal{P}(X)$.

If $\mathcal{F}$ is any (proper) filter on $\mathcal{P}(X)$ such that $\mathcal{U} \subseteq \mathcal{F}$, then $\mathcal{U} = \mathcal{F}$.

Proof. Assume there were some $Y \in \mathcal{F} - \mathcal{U}$. Then $X - Y \in \mathcal{U}$ by property 3 of an ultrafilter, and hence $X - Y \in \mathcal{F}$. But then $Y \cap (X - Y) \in \mathcal{F}$, by property 1 of a filter. So $\emptyset \in \mathcal{F}$, contradicting the fact that $\mathcal{F}$ is a proper filter. $\square$

Lemma 3.3.2.2. Let $\mathcal{F}$ be a (proper) filter on $\mathcal{P}(X)$, and let $W \in \mathcal{P}(X)$.

Then either W meets everything in $\mathcal{F}$ or its complement $X - W$ meets everything in $\mathcal{F}$.

More explicitly:

$$\forall_{Y \in \mathcal{F}}(Y \cap W \neq \emptyset) \qquad \text{or} \qquad \forall_{Y \in \mathcal{F}}(Y \cap (X - W) \neq \emptyset)$$

Proof. Assume W does not meet everything in $\mathcal{F}$. Then there is $Z \in \mathcal{F}$ such that $W \cap Z = \emptyset$. Then $Z \subseteq (X - W)$, so $(X - W) \in \mathcal{F}$. $\square$

Remark. The above proof shows that if W does not meet everything in $\mathcal{F}$, then its complement $(W - X)$ is in $\mathcal{F}$. However, W does not have to be in $\mathcal{F}$ to meet all its sets, and it is possible for both W and $X - W$ to meet everything in $\mathcal{F}$. Consider, for example, the cofinite filter on $\mathbb{N}$ and the sets of even and odd numbers. $\triangleleft$

Proposition 3.3.2.3. Let $\mathcal{U} \subseteq \mathcal{P}(X)$ be a (proper) filter, with X non-empty, such that for any other (proper) filter $\mathcal{F} \subseteq \mathcal{P}(X)$, if $\mathcal{U} \subseteq \mathcal{F}$ then $\mathcal{U} = \mathcal{F}$.

Then $\mathcal{U}$ is an ultrafilter.

Proof. If $\mathcal{U}$ were not an ultrafilter, then there would be some set $W \in \mathcal{P}(X)$ such that neither W nor $X - W$ is in $\mathcal{U}$.

By Lemma 3.3.2.2, either W or $X - W$ meets every set in $\mathcal{U}$. Assume without loss of generality that it is W which meets every set in $\mathcal{U}$.

Construct sets

$$G = \{W \cap Z : Z \in \mathcal{U}\}.$$

$$\mathcal{F} = \{Z \in \mathcal{P}(X) : Z \supseteq Y, \text{ some } Y \in G\}$$

(Note that $G \subseteq \mathcal{F}$, and, in particular, since $X \in \mathcal{U}$ then $W \in G \subseteq \mathcal{F}$).

It is not difficult to show G is closed under intersection and $\mathcal{F}$ is closed under taking supersets. It is shown $\mathcal{F}$ is also closed under intersection.

Let $A, B \in \mathcal{F}$. Then $A \supseteq J$ and $B \supseteq K$ for some $J, K \in G$.

Both $J = W \cap Y$ and $K = W \cap Z$, for some $Y, Z \in \mathcal{U}$. Then $J \cap K = (W \cap Y) \cap (W \cap Z) = W \cap (Y \cap Z)$ and $Y \cap Z \in \mathcal{U}$. So $J \cap K \in G$. Since $A \cap B \supseteq J \cap K$, then $A \cap B \in \mathcal{F}$ also.

By assumption, $W \cap Z \neq \emptyset$ for any $Z \in U$. As $\emptyset \notin \mathcal{U}$, also $\emptyset \notin G$. Thus $\emptyset \notin \mathcal{F}$.

Also, for any $Z \in \mathcal{U}$, since $Z \supseteq Z \cap W \in \mathcal{F}$, then also $Z \in \mathcal{F}$. So $\mathcal{U} \subseteq \mathcal{F}$.

Then $\mathcal{F}$ is a (proper) filter on X strictly containing $\mathcal{U}$, contradicting the fact that $\mathcal{U}$ is maximal.

Hence $\mathcal{U}$ is an ultrafilter. $\square$

Corollary 3.3.2.4. Let X be non-empty, and let $\mathcal{U} \subseteq \mathcal{P}(X)$. Then $\mathcal{U}$ is an ultrafilter if and only if it is maximal as a filter on $\mathcal{P}(X)$. $\lhd$

3.3.3 Non-principal Ultrafilters

Proving the existence of non-principal ultrafilters requires something stronger than the Zermelo-Fraenkal axioms. The axiom of choice is sufficient, but existence of non-principal ultrafilters is strictly weaker. A common existence theorem (weaker than AC) is the Boolean Prime Ideal theorem (BPI). See [DP02] for more details.

The version of BPI Theorem for ultrafilters on a powerset algebra will suffice. The proof is given below using Zorns lemma, and adapted from [BS71, p15]:

Theorem 3.3.3.1. Every filter on $\mathcal{P}(X)$ can be extended to an ultrafilter on $\mathcal{P}(X)$.

Proof. Let $\mathcal{F}$ be a filter on $\mathcal{P}(X)$. Let $\mathbb{F}$ be the set of all proper filters containing $\mathcal{F}$, ordered by inclusion. It is obviously not empty, since $\mathcal{F} \in \mathbb{F}$.

It is shown every chain (linearly ordered sub-poset) $\mathbb{C}$ in $\mathbb{F}$ has an upper-bound. By Zorn's lemma, it then follows that $\mathbb{F}$ contains a maximal element $\mathcal{U}$ (containing $\mathcal{F}$), which is hence an ultrafilter by Theorem 3.3.2.4.

Write $\mathbb{C}$ as $(C_i)_{i \in I}$ where I is a linearly ordered set, and $i \leq j \Rightarrow C_i \subseteq C_j$. Let $\mathcal{C} = \cup\mathbb{C} = \cup_{i \in I} C_i$.

This is a (proper) filter:

- If $Y, Z \in \mathcal{C}$, then $Y \in C_j$, and $Z \in C_k$ for some j and k. Both C_j and C_k are elements of a chain. Without loss of generality, assume $C_j \subseteq C_k$. Then $Y, Z \in C_k$ and so $Y \cap Z \in C_k$.

- Similarly, if $Y \in \mathcal{C}$ then $Y \in C_j$ for some j, so if $Z \supseteq Y$ then $Z \in C_j \subseteq \mathcal{C}$.
- If $\emptyset \in \mathcal{C}$ then $\emptyset \in C_j$ for some j which is not possible since all the C_j are proper filters.

Hence $\mathcal{C}$ is a filter which is an upper bound for $(C_i)_{i\in I}$. Thus, every chain has an upper bound and so $\mathbb{F}$ has a maximal element which is an ultrafilter. □

3.4 Partition Condition

The following is stated in [Lei13] but is originally formulated and proven in [GH70].

Definition 3.4.0.1. A subset $\mathcal{U} \subseteq \mathcal{P}(X)$ satisfies the n**-partition condition** if, for any partition of X into n disjoint (possibly empty) subsets covering X:

$$\begin{aligned} &X = X_1 \cup \cdots \cup X_n, \\ &X_i \cap X_j = \emptyset, \text{ all } i \neq j, \end{aligned}$$

$\mathcal{U}$ contains the subset X_i for exactly one of the $i \leq n$. ◁

Lemma 3.4.0.2. Let X be a non-empty set and $\mathcal{U} \subseteq \mathcal{P}(X)$ a subset of $\mathcal{P}(X)$.

Then the following are equivalent:

1. $\mathcal{U}$ is an ultrafilter.
2. $\mathcal{U}$ satisfies the 3-partition condition.
3. $\mathcal{U}$ satisfies the n-partition condition for any $n \geq 3$.

Proof. Proofs are available in [Lei13], and in the original paper [GH70]. □

3.5 Ultrafilter Monad

3.5.1 Ultrafilter Functor

The functor $U : \mathbf{Set} \to \mathbf{Set}: X \to U(X)$ sends a set X to the set $U(X)$ of all ultrafilters on $\mathcal{P}(X)$. It sends a morphism $f : X \to Y$ to the morphism

$$U(f) : U(X) \to U(Y) : \mathcal{U} \mapsto f_*\mathcal{U} = \{Z \subseteq Y : f^{-1}Y \in \mathcal{U}\}.$$

3.5.2 Monad Structure

The next few results are due to [Bör87]:

Theorem 3.5.2.1.

U preserves finite coproducts. $\lhd$

It is terminal as an endofunctor preserving finite coproducts:

Theorem 3.5.2.2. If $F : \mathbf{Set} \to \mathbf{Set}$ is an endofunctor preserving finite coproducts, then there exists a unique natural transformation $\kappa : F \to U$. $\lhd$

It is the functor of a monad:

Lemma 3.5.2.3. There exist unique natural transformations $\eta : 1_{\mathbf{Set}} \to U$ and $\mu : U \circ U \to U$. $\lhd$

Corollary 3.5.2.4. (U, η, μ) is a monad. $\lhd$

Furthermore, it is terminal as a finite coproduct preserving monad:

Corollary 3.5.2.5. If (V, ϵ, ν) is another monad over **Set** such that V preserves finite coproducts, then there is a unique monad morphism $\kappa : (V, \epsilon, \nu) \to (U, \eta, \mu)$. $\lhd$

The following corollary will be used in §8.8 to show that a certain codensity monad is equal to the ultrafilter monad.

Corollary 3.5.2.6. There is a unique monad structure on the functor U.

Proof. By lemma 3.5.2.3, there exist unique natural transformations $\eta : 1_{\mathbf{Set}} \to U$ and $\mu : U \circ U \to U$, so they must form the unit and multiplication of the monad. $\square$

Corollary 3.5.2.7. If (V, ϵ, ν) is a monad and $U \cong V$ is a natural isomorphism of functors, then

$$(U, \eta, \mu) \cong (V, \epsilon, \nu).$$

Proof. Let $\alpha : U \cong V$ be the natural isomorphism with inverse α^{-1}.

There is natural isomorphism

$$V(\alpha) \circ \alpha_U : U \circ U \Rightarrow V \circ V$$

given at $X \in \mathbf{Set}$ by

$$(V(\alpha) \circ \alpha_U)_X = V(\alpha_X) \circ \alpha_{U(X)}.$$

It has an inverse

$$(V(\alpha) \circ \alpha_U)^{-1} = \alpha_U^{-1} \circ V(\alpha)^{-1} : V \circ V \Rightarrow U \circ U$$

given at $X \in \mathbf{Set}$ by

$$\left((V(\alpha) \circ \alpha_U)^{-1}\right)_X = \alpha_{U(X)}^{-1} \circ V\left(\alpha_X^{-1}\right).$$

By Lemma 3.5.2.3, there exist unique natural transformations $\eta : 1_{\mathbf{Set}} \to U$ and $\mu : U \circ U \to U$.

Hence, since $\alpha \circ \eta : 1_{\mathbf{Set}} \to V$ and $\alpha \circ \mu \circ (V(\alpha) \circ \alpha_U)^{-1} : V \circ V \to V$ are natural transformations between the relevant functors then

$$\epsilon = \alpha \circ \eta$$

and

$$\nu = \alpha \circ \mu \circ (V(\alpha) \circ \alpha_U)^{-1}.$$

So α is a monad morphism and hence the two monads are isomorphic. $\square$

Chapter 4

Ultraproducts

4.1 Ultraproduct: Standard Definition

The construction below mostly follows [CK73, 167-170], but is standard and can be found in many other textbooks which concern ultraproducts (such as [CN70, 269; BS71, 88-89]).

An ultraproduct is a special case of a reduced product, where the filter is an ultrafilter.

4.2 Underlying Set of the Reduced Product

Let I be an indexing set (usually infinite). Let $\langle A_i \rangle_{i \in I}$ be a family of non-empty sets indexed by I. Let $\mathcal{F}$ be a filter on $\mathcal{P}(I)$.

Define

$$C = \prod_{i \in I} A_i$$

A tuple $a \in C$ can be written $a = \langle A_i \rangle_{i \in I}$ where each $a_i \in A_i$.

For two tuples $a, b \in C$ given by $a = \langle A_i \rangle_{i \in I}$ and $b = \langle b_i \rangle_{i \in I}$, define a relation:

$a \sim b$ iff $\{i \in I : a_i = b_i\} \in \mathcal{F}$

Lemma 4.2.0.1. If $\mathcal{F}$ is a non-empty filter, the above relation is an equivalence relation on C

Proof. If $a \sim b$ then obviously $b \sim a$.

Reflexivity follows from the fact that $I \in \mathcal{F}$ since any non-empty filter always contains the whole set.

For transitivity, assume $a \sim b$ and $b \sim c$. Then from the properties of a filter we have

$$\begin{aligned}
&\{i \in I : a_i = b_i\} \in \mathcal{F} \text{ and } \{i \in I : b_i = c_i\} \in \mathcal{F} \\
&\{i \in I : a_i = b_i\} \cap \{i \in I : b_i = c_i\} \in \mathcal{F} \\
&\{i \in I : a_i = c_i\} \supseteq \{i \in I : a_i = b_i \text{ and } b_i = c_i\} \in \mathcal{F} \\
&\{i \in I : a_i = c_i\} \in \mathcal{F}
\end{aligned}$$

□

Now the relation $\sim$ defines an equivalence relation on C. An element of $C/\sim$ will be written as $[\langle A_i \rangle_{i \in I}]$, or simply $[a]$, to denote the equivalence class of $a \in C$.

Definition 4.2.0.2. $C/\sim$ is referred to as the **reduced product** of $\langle A_i \rangle_{i \in I}$ modulo the filter $\mathcal{F}$. It may be written as

$$\prod_{\mathcal{F}} A_i$$

If the filter is an ultrafilter, then $\prod_{\mathcal{F}} A_i$ is referred to as an **ultraproduct**.

◁

Remark. The above notation is somewhat ambiguous, but will be precise enough for our current purposes. ◁

4.2.1 Model-Theoretic Structure of the Reduced Product

Now, for the model-theoretic structure, fix a language L with relation symbols R^j, function symbols f^j, and constants c^j interpreted in A_i by the relations, functions and constants R_i^j, f_i^j and c_i^j respectively. Their interpretation in $\prod_{\mathcal{F}} A_i$ is denoted by $R_{\mathcal{F}}^j$, f_F^j and $c_{\mathcal{F}}^j$ respectively. For simplicity, the superscript js will be omitted in the following, and in general both subscripts and superscripts will be omitted.

Definition 4.2.1.1. Let $x_i^1, \cdots x_i^n \in A_i$.

Each constant c is interpreted in $\prod_{\mathcal{F}} A_i$ by

$$c_{\mathcal{F}} = [\langle c_i \rangle_{i \in I}]$$

Each n-ary function is interpreted by

$$f_{\mathcal{F}}([\langle x_i^1 \rangle_{i \in I}], \cdots, [\langle x_i^n \rangle_{i \in I}]) = [\langle f_i(x_i^1, \cdots, x_i^n) \rangle_{i \in I}]$$

Each n-ary relation symbol is interpreted by

$$R_{\mathcal{F}}([\langle x_i^1 \rangle_{i\in I}], \cdots, [\langle x_i^n \rangle_{i\in I}]) \qquad \text{iff} \qquad \{i \in I : R_i(x_i^1, \cdots, x_i^n)\} \in \mathcal{F}$$

$\triangleleft$

It remains to prove that the above definitions are consistent, i.e., that the values of constants and functions, and the truth of relations depend only on the equivalence classes and not on their representatives. It is sufficient to show this only for functions and relations (the proof for constants then follows by taking them as 0-arity functions).

Theorem 4.2.1.2. For $x^1, \cdots, x^n, y^1, \cdots y^n \in \prod_{i\in I} A_i$, let $\langle x_i^j \rangle_{i\in I} \sim \langle x_i^j \rangle_{i\in I}$ for all $j \leq n$.

Then

$$f_{\mathcal{F}}([\langle x_i^1 \rangle_{i\in I}], \cdots, [\langle x_i^n \rangle_{i\in I}]) = f_{\mathcal{F}}([\langle y_i^1 \rangle_{i\in I}], \cdots, [\langle y_i^n \rangle_{i\in I}])$$

and

$$R_{\mathcal{F}}([\langle x_i^1 \rangle_{i\in I}], \cdots, [\langle x_i^n \rangle_{i\in I}]) \qquad \text{iff} \qquad R_{\mathcal{F}}([\langle y_i^1 \rangle_{i\in I}], \cdots, [\langle y_i^n \rangle_{i\in I}])$$

Proof. Since $x^j \sim y^j$, then $\{i \in I : x_i^j = y_i^j\} \in \mathcal{F}$ for each $j \leq n$, and hence also the finite intersection across all $j \leq n$ is in $\mathcal{F}$. So

$$\{i \in I : \text{ for all } j \leq n, x_i^j = y_i^j\} \in \mathcal{F}$$

Hence

$$\begin{aligned} f_{\mathcal{F}}([\langle x_i^1 \rangle_{i\in I}], \cdots, [\langle x_i^n \rangle_{i\in I}]) &= [\langle f_i(x_i^1, \cdots, x_i^n) \rangle_{i\in I}] \\ &= [\langle f_i(y_i^1, \cdots, y_i^n) \rangle_{i\in I}] \\ &= f_{\mathcal{F}}([\langle y_i^1 \rangle_{i\in I}], \cdots, [\langle y_i^n \rangle_{i\in I}]) \end{aligned}$$

and

$$\begin{aligned} R_{\mathcal{F}}([\langle x_i^1 \rangle_{i\in I}], \cdots, [\langle x_i^n \rangle_{i\in I}]) \qquad &\text{iff} \qquad \{i \in I : R_i(x_i^1, \cdots, x_i^n)\} \in \mathcal{F} \\ &\text{iff} \qquad \{i \in I : R_i(y_i^1, \cdots, y_i^n)\} \in \mathcal{F} \\ &\text{iff} \qquad R_{\mathcal{F}}([\langle y_i^1 \rangle_{i\in I}], \cdots, [\langle y_i^n \rangle_{i\in I}]) \end{aligned}$$

$\square$

Remark. Although the definition is well-defined for any non-empty proper filter, an ultrafilter is required for Łoś's theorem. $\triangleleft$

4.3 Ultraproduct: Category-Theoretic Definition

The first known source that defines ultraproducts category-theoretically as a filtered colimit of products is [Okh66].

Let $\mathbb{C}$ be a category with small products and filtered colimits.

Let $(A_i)_{i \in I}$ be a family of objects of the category, indexed by the set I, and $\mathcal{U} \subseteq \mathcal{P}(I)$ be an ultrafilter.

$\mathcal{U}$ is a poset ordered by subset inclusion. Consider the inverse of this poset, (ordered by reverse inclusion), then it can be viewed as a category, with morphisms given by maps:

$$Y \to X \qquad \text{iff} \qquad Y \supseteq X.$$

Define a functor

$$\mathcal{U} \xrightarrow{P} \mathbb{C}$$

$$\begin{tikzcd}
Y \arrow[d, "\supseteq"'] \arrow[r, phantom, ""] & \prod_{y \in Y} A_y \arrow[d, "\pi_X^Y"] \\
X & \prod_{x \in X} A_x
\end{tikzcd}
\qquad (\longmapsto)$$

Where the π morphisms are the product morphisms induced by the projections $\pi_x : (\prod_{y \in Y} A_y) \to A_x$ for $x \in X$. In the case of sets, these are the maps forgetting those coordinates that do not occur in X.

Definition 4.3.0.1. The filtered colimit

$$\varinjlim P$$

is the **category-theoretic ultraproduct** of $\langle A_i \rangle_{i \in I}$ with respect to the ultrafilter $\mathcal{U}$. $\lhd$

This colimit may be written

$$\varinjlim P = \varinjlim_{X \in \mathcal{U}} \prod_{x \in X} A_x.$$

A version of the following theorem may be found in e.g., [Ekl77].

Theorem 4.3.0.2. If $\mathbb{C}$ is the category of **sets**, then

$$\varinjlim P \cong \left(\sum_{X \in \mathcal{U}} \prod_{x \in X} A_x \right) \Big/ \sim$$

Where Σ denotes the coproduct (disjoint union) and the equivalence relation $\sim$ is given by

$$\langle a_x \rangle_{x \in X} \sim \langle b_x \rangle_{x \in Y} \qquad \text{iff} \qquad \{x \in X \cap Y : a_x = b_x\} \in \mathcal{U}$$

For ease of notation, write

$$\prod_{\mathcal{U}} A_x := \left(\sum_{X \in \mathcal{U}} \prod_{x \in X} A_x \right) \Big/ \sim$$

The morphisms of the colimiting cone are given by

$$\amalg_X : \prod_{x \in X} A_x \to \prod_{\mathcal{U}} A_x : \langle a_x \rangle_{x \in X} \mapsto [\langle a_x \rangle_{x \in X}],$$

where $[\langle a_x \rangle_{x \in X}]$ is the equivalence class of an element in the subset $\prod_{x \in X} A_x$ of the disjoint union.

Proof. (Also see the diagrams below the proof).

If for every large set $X \in \mathcal{U}$ there some $x \in X$ such that $A_x = \emptyset$, then all of the products are empty, and so the isomorphism is trivial. Hence, assume that there is some large set $J \in \mathcal{U}$ such that all of the A_x are non-empty on J.

It must be shown that $\sim$ is an equivalence relation. Both symmetry and reflexivity are trivial. Let $\langle a_x \rangle_{x \in X}, \langle b_x \rangle_{x \in Y}, \langle c_x \rangle_{x \in Z}$ be three families such that

$$\langle a_x \rangle_{x \in X} \sim \langle b_x \rangle_{x \in Y} \quad \text{and} \quad \langle b_x \rangle_{x \in X} \sim \langle c_x \rangle_{x \in Y}.$$

Then

$$\{x \in X \cap Y : a_x = b_x\} \in \mathcal{U} \quad \text{and} \quad \{x \in Y \cap Z : b_x = c_x\} \in \mathcal{U}$$

so their intersection is also in the ultrafilter, and hence also the set

$$\{x \in X \cap Z : a_x = c_x\} \supseteq \{x \in X \cap Y \cap Z : a_x = b_x = c_x\} \in \mathcal{U}.$$

Remark. The above argument applies even if $\mathcal{U}$ is merely a **filter**. $\lhd$

Now it is shown that the above construction is indeed the colimit of the given cone. Let W be any other object with a cone $w_X : \prod_{x\in X} A_x \to W$ which commutes with the π morphisms. Then it remains to show that there is a unique morphism $h : \prod_{\mathcal{U}} A_x \to W$ such that $h\amalg_X = w_X$.

Assume this h exists. Let $\bar{a} \in \prod_{\mathcal{U}} A_x$. Then

$$\bar{a} = [\langle a_x\rangle_{x\in X}] = \amalg_X(\langle a_x\rangle_{x\in X}) \text{ for some (non-empty) } X \in \mathcal{U}.$$

Since $h\amalg_X = w_X$, then

$$h([\langle a_x\rangle_{x\in X}]) = h(\amalg_X(\langle a_x\rangle_{x\in X})) = w_X(\langle a_x\rangle_{x\in X}).$$

Hence h is uniquely determined.

Now it must be shown that if h is defined as above, its image is independent of the choice of representative of the equivalence class.

Assume $\langle a_x\rangle_{x\in X} \sim \langle b_x\rangle_{x\in Y}$. Then there is $Z \in \mathcal{U}$, with $Z \subseteq X \cap Y$, such that $a_x = b_x$ for $x \in Z$, and so $\langle a_x\rangle_{x\in Z} = \langle b_x\rangle_{x\in Z}$.

Hence,

$$\begin{aligned}
h([\langle a_x\rangle_{x\in X}]) &= h\amalg_X(\langle a_x\rangle_{x\in X}) \\
&= w_X(\langle a_x\rangle_{x\in X}) = w_Z(\pi^X_Z\langle a_x\rangle_{x\in X}) \\
&= w_Z(\langle a_x\rangle_{x\in Z}) = w_Z(\langle b_x\rangle_{x\in Z}) \\
&= w_Z(\pi^Y_Z\langle b_x\rangle_{x\in Y}) = w_Y(\langle b_x\rangle_{x\in Y}) \\
&= h\amalg_Y(\langle b_x\rangle_{x\in Y}) = h([\langle b_x\rangle_{x\in Y}])
\end{aligned}$$

□

$$\begin{tikzcd}
& W & & \\
& & \prod_{\mathcal{U}} A_x \arrow[ul, dotted, "h"'] & \\
\prod_{x\in X} A_x \arrow[uur, bend left, "w_X"] \arrow[urr, "\amalg_X"] \arrow[dr, "\pi^X_Z"'] & & & \prod_{x\in Y} A_x \arrow[uull, bend right, "w_Y"'] \arrow[ul, "\amalg_Y"'] \arrow[dll, "\pi^Y_Z"] \\
& \prod_{x\in Z} A_x \arrow[uuu, bend left, "w_Z"] \arrow[uur, "\amalg_Z"] & &
\end{tikzcd}$$

$$\begin{tikzcd}
 & w & & \\
 & & {[\langle a_x\rangle_{x\in Z}]} \arrow[ul, dotted, "h"'] & \\
\langle a_x\rangle_{x\in X} \arrow[uur, bend left, "w_X"] \arrow[urr, "\rotatebox[origin=c]{180}{π}_X"'] \arrow[dr, "\pi^X_Z"'] & & & \langle b_x\rangle_{x\in Y} \arrow[uull, bend right, "w_Y"'] \arrow[ul, "\rotatebox[origin=c]{180}{π}_Y"] \arrow[dll, "\pi^Y_Z"] \\
 & \langle a_x\rangle_{x\in Z} \arrow[uuu, bend left, "w_Z"] \arrow[uur, "\rotatebox[origin=c]{180}{π}_Z"'] & &
\end{tikzcd}$$

The following proposition shows that for sets, the ultraproduct can be simplified to a quotient of a product.

Proposition 4.3.0.3. Let $J \subseteq I$ be defined as those indices on which the sets are non-empty:

$$J = \{x \in I : A_x \neq \emptyset\}$$

Case 1: $J \notin \mathcal{U}$. Then

$$\left(\sum_{X\in\mathcal{U}}\prod_{x\in X} A_x\right)\Big/\sim \qquad = \qquad \emptyset$$

Case 2: $J \in \mathcal{U}$.

Define the equivalence relation $\sim_J$ on $\prod_{x\in J} A_x$ by

$$\langle a_x\rangle_{x\in J} \sim_J \langle b_x\rangle_{x\in J} \qquad \text{iff} \qquad \{x \in J : a_x = b_x\} \in \mathcal{U}$$

Then there is a bijection

$$\left(\prod_{x\in J} A_x\right)\Big/\sim_J \qquad \cong \qquad \left(\sum_{X\in\mathcal{U}}\prod_{x\in X} A_x\right)\Big/\sim$$

given by

$$[\langle a_x\rangle_{x\in J}] \mapsto [\langle a_x\rangle_{x\in J}].$$

Proof. If $X \in \mathcal{U}$ and there is some $x \in X$ with $A_x = \emptyset$, then $\prod_{x\in X} A_x = \emptyset$.

If $J \notin \mathcal{U}$, then $X \in \mathcal{U}$ means there is $x \in X$ with $A_x = \emptyset$ (since otherwise $X \subseteq J$ and so $J \in \mathcal{U}$). Hence $\prod_{x\in X} A_x = \emptyset$ so $\sum_{X\in\mathcal{U}}\prod_{x\in X} A_x = \emptyset$.

Assume $J \in \mathcal{U}$.

Let $\mathcal{V} \subseteq \mathcal{U}$ with $\mathcal{V} = \{X \in \mathcal{U} : A_x \neq \emptyset \text{ all } x \in X\}$. Hence $\mathcal{V} = \mathcal{P}(J) \cap \mathcal{U}$.

Note that $\mathcal{V}$ is closed under intersection and superset).

Define a relation $\sim_{\mathcal{V}}$ on $X, Y \in \mathcal{V}$ by

$$\langle a_x \rangle_{x \in X} \sim_{\mathcal{V}} \langle b_x \rangle_{x \in Y} \qquad \text{iff} \qquad \{x \in X \cap Y : a_x = b_x\} \in \mathcal{V}$$

Symmetry and reflexivity of this relation follow from symmetry and reflexivity of equality.

Transitivity follows from the fact that $\{x \in X \cap Y \mid a_x = b_x\} \in \mathcal{V}$ and $\{x \in Y \cap Z \mid b_x = c_x\} \in \mathcal{V}$ implies that their intersection $\{x \in X \cap Y \cap Z \mid a_x = b_x = c_x\} \in \mathcal{V}$ and so the superset $\{x \in X \cap Z \mid a_x = c_x\}$ is also.

Now, consider $X \in \mathcal{U}$ with some $x \in X$ such that $A_x = \emptyset$. Then the product

$$\prod_{x \in X} A_x = \emptyset.$$

Hence, in the disjoint union, the products over such X may be ignored. Hence, the disjoint union may be written as

$$\sum_{X \in \mathcal{U}} \prod_{x \in X} A_x \quad = \quad \sum_{X \in \mathcal{V}} \prod_{x \in X} A_x.$$

And so the quotients are also bijective

$$\left(\sum_{X \in \mathcal{U}} \prod_{x \in X} A_x \right) \Big/ \sim \qquad \cong \qquad \left(\sum_{X \in \mathcal{V}} \prod_{x \in X} A_x \right) \Big/ \sim_{\mathcal{V}}$$

Hence it remains only to show there is a bijection:

$$\left(\prod_{x \in J} A_x \right) \Big/ \sim_J \qquad \cong \qquad \left(\sum_{X \in \mathcal{V}} \prod_{x \in X} A_x \right) \Big/ \sim_{\mathcal{V}}$$

It must be shown the map is well-defined (not dependent of choice of representative for the equivalence class), and that there is an inverse.

Define functions

$$f : \left(\prod_{x \in J} A_x \right) \Big/ \sim_J \to \left(\sum_{X \in \mathcal{V}} \prod_{x \in X} A_x \right) \Big/ \sim_{\mathcal{V}} : [\langle a_x \rangle_{x \in J}]_J \mapsto [\langle a_x \rangle_{x \in J}]_{\mathcal{V}}$$

and

$$g : \left(\sum_{X \in \mathcal{V}} \prod_{x \in X} A_x \right) \Big/ \sim_{\mathcal{V}} \to \left(\prod_{x \in J} A_x \right) \Big/ \sim_J : [\langle a_x \rangle_{x \in J}]_{\mathcal{V}} \mapsto [\langle a_x \rangle_{x \in J}]_J.$$

It must be shown that these functions are well-defined.

Let $\langle a_x \rangle_{x\in J} \sim_J \langle b_x \rangle_{x\in J}$. Then since $J \in \mathcal{V}$, it holds that $\langle a_x \rangle_{x\in J} \sim_{\mathcal{V}} \langle b_x \rangle_{x\in J}$, so the map f is well-defined.

For the inverse g: It must be shown that for any $X \in \mathcal{V}$ and any $\langle a_x \rangle_{x\in X}$, the sequence is equivalent to one indexed by J, i.e., $\langle a_x \rangle_{x\in X} \sim_{\mathcal{V}} \langle a'_x \rangle_{x\in J}$.

Each of the sets A_x for $x \in J - X$ is non-empty, and so it is possible to choose an element c_x from each of them. Complete $\langle a_x \rangle_{x\in X}$ to $\langle a'_x \rangle_{x\in J}$ by letting $a'_x = a_x$ for $x \in X$ and $a'_x = c_x$ for $x \in J - X$. Since the set $J - X$ on which the c_x are chosen is small, the resulting equivalence class is independent of the choices of c_x.

This map is independent of choice of representative: Let $\langle a_x \rangle_{x\in X} \sim_{\mathcal{V}} \langle b_x \rangle_{x\in Y}$. Then $X, Y \in \mathcal{V}$ and so $\langle a_x \rangle_{x\in X\cap Y} = \langle b_x \rangle_{x\in X\cap Y}$.

Using the above procedure, complete $\langle a_x \rangle_{x\in X\cap Y}$ to $\langle a'_x \rangle_{x\in J}$. Then since $J - X$, $J - Y$ and $J - X \cap Y$ are all small, the map g acts the same regardless of whether it is $\langle a_x \rangle_{x\in X}$, $\langle b_x \rangle_{x\in X}$ or $\langle a_x \rangle_{x\in X\cap Y}$ that is completed.

That f and g are inverse to each other follows from their definitions. □

Corollary 4.3.0.4. If all of the sets A_x are non-empty, then the construction simplifies to

$$\varinjlim P \cong \prod_{\mathcal{U}} A_x \cong \left(\prod_{x\in I} A_x \right) \Big/ \sim$$

with $\langle a_x \rangle_{x\in I} \sim \langle b_x \rangle_{x\in I}$ iff $\{x \in I : a_x = b_x\} \in \mathcal{U}$. ◁

This is indeed the way the ultraproduct is usually defined in model theory.

Remark. One needs to be careful about the notation used for sequences when the sets are allowed to be empty. If $A_x = \emptyset$ for some $x \in X$, then $\langle a_x \rangle_{x\in X}$ does not make sense, and so one cannot speak about $[\langle a_x \rangle_{x\in X}]$. By convention, we will define $[\langle a_x \rangle_{x\in X}] = [\langle a_x \rangle_{x\in X\cap J}]$ where J is, as above, the set of indices on which the A_i are non-empty. ◁

4.4 Problems With the Category Theoretic Notions of Ultraproducts

4.4.1 Overview

This section (§4.4) is based on the paper [BN87].

Not every category has small products and filtered colimits. Hence, it is not possible to form the category-theoretic ultraproduct. This is the case even in some concrete categories, where it possible to form an ultraproduct using sets.

For example, the category of fields does not have all small products, but the model-theoretic ultraproduct may be formed as usual.

In [BN87], and in this section, it is further shown that there is a category in which small products and directed colimits exist, and so category-theoretic ultraproducts may be formed. However, the ultraproduct as calculated in this category is different from the model-theoretic ultraproduct (and hence also from the ultraproduct as calculated in the category of relational structures).

At first glance, this may seem to contradict the result that Łoś's Theorem holds in every category (§7.1). However, this apparent contradiction is resolved by realising §7.1 only proves that trees represent all first-order formulae in categories of **relational structures**. In other categories, the theorem still holds with regards to whatever trees may be definable, but these trees don't necessarily represent all first-order formulae.

Further, it is argued here that the existence of categories (of algebraic structures) in which ultraproducts do not correspond to the model-theoretic version is not a big problem, since the usual ultraproduct can still be calculated category-theoretically simply by extending to an appropriate category (namely, a variety of algebras) and indicates only that one needs to be careful not to assume that the ultraproducts will correspond in general.

Remark. In this section, it does not matter whether or not $\mathbb{N}$ contains 0, as long as it is consistent. This section assumes $0 \notin \mathbb{N}$. $\lhd$

4.4.2 Category With Different Ultraproducts

A family of structures will be defined. In this section, they are referred to as "ultraproduct-problematic structures", or just "problematic structures" for short.

Define a countable language $\langle Z, P_1, P_2 \ldots, S_1, S_2 \ldots \rangle$ consisting of a unary relation symbol Z, unary relation symbols P_i for $i \in \mathbb{N}$, and unary relation symbols S_i for $i \in \mathbb{N}$.

Define a countable system of axioms consisting of the statements

$$\forall_x Z(x) \to \forall_x [P_i(x) \leftrightarrow \neg S_i(x)]$$

for each $i \in \mathbb{N}$.

Now consider the category $\mathbb{C}$ of models of these axioms (as a subcategory of the category of relational structures). It is described here for clarity.

- An object X in this category is a set X and a unary relation (subset of X) for each relation such that the axioms hold.

- A morphism $f : X \to Y$ is a set-map preserving the relations in the sense that for each relation symbol R, if $R^X(x)$ in X then $R^Y(f(x))$ in Y. (In subset notation, $f[R^X] \subseteq R^Y$).

Fix a (non-principal) ultrafilter $\mathcal{U} \subseteq \mathcal{P}(\mathbb{N})$.

Let $\langle A_n \rangle_{n\in\mathbb{N}}$ be a family of such structures defined as follows. For each $n \in \mathbb{N}$,

- the underlying set is a singleton $A_n = \{a\}$,
- for each $i \leq n$, $A_n \models P_i(a) \wedge \neg S_i(a)$,
- for each $i > n$, $A_n \models \neg P_i(a) \wedge S_i(a)$,
- $A_n \models Z(a)$,

Remark. Note, in particular, this means that $A_n \models \forall_x Z(x)$ and similarly for P_i, $\neg P_i$, S_i and $\neg S_i$. $\lhd$

Consider the model-theoretic ultraproduct

$$\prod_{\mathcal{U}} A_n.$$

By Łoś's Theorem,

- The underlying set is a singleton $\prod_{\mathcal{U}} A_n = \{a\}$
- For each $i \in \mathbb{N}$, $\prod_{\mathcal{U}} A_n \models P_i(a) \wedge \neg S_i(a)$.
- $\prod_{\mathcal{U}} A_n \models Z(a)$.

Now consider the category-theoretic ultraproduct

$$\varinjlim_{H\in\mathcal{U}} \prod_{n\in H} A_n.$$

Assume, for the moment, that it exists, i.e., that the relevant products and directed colimit exist.

Consider a specific product $\prod_{n\in H} A_n$. This has projection maps $\pi_n : \prod_{n\in H} A_n \to A_n$ for each $n \in \mathbb{N}$. Since H is infinite, then for any $i \in \mathbb{N}$, there is some $n > i$ such that $n \in H$.

What this implies is that for any i, there is a large enough n so that $A_n \models \neg S_i(a)$. But since the projection map π_n must preserve relations, and since every x in the product is mapped to $a \in A_n$, then no x in the product can satisfy $S_i(x)$. (We do not yet know whether the product is a singleton). Hence, for all $i \in \mathbb{N}$,

$$\prod_{n \in H} A_n \models \forall_x \neg S_i(x).$$

Furthermore, consider the least $n \in H$. Then $A_n \models \neg P_i(a)$ for all $i > n$. So $\forall_x \neg P_i(x)$ holds for all $i > n$ in the product.

Thus, for example, $\forall_x \neg P_{n+1}(x) \wedge \neg S_{n+1}(x)$, in the product. So by (the contrapositive of) the axiom,

$$\forall_x Z(x) \to \forall_x [P_{n+1}(x) \leftrightarrow \neg S_{n+1}(x)]$$

it must hold that $\neg \forall_x Z(x)$ in the product.

Now, in fact, such products do exist and are given as follows. Let $H \in \mathcal{U}$, and let m be the least element of H. Then:

- The underlying set is a singleton $\prod_{n \in H} A_n = \{a\}$.
- For each $i \in \mathbb{N}$, $\prod_{n \in H} A_n \models \neg S_i(a)$.
- For each $i \leq m$, $\prod_{n \in H} A_n \models P_i(a)$.
- For each $i > m$, $\prod_{n \in H} A_n \models \neg P_i(a)$.
- $\prod_{n \in H} A_n \models \neg Z(a)$

It is routine to show that this object, along with projections given by the unique maps into each A_n satisfies the universal property necessary for the product.

Now, consider the directed colimit.

$$\varinjlim_{H \in \mathcal{U}} \prod_{n \in H} A_n.$$

It is given as follows:

- The underlying set is a singleton $\varinjlim_{H \in \mathcal{U}} \prod_{n \in H} A_n = \{a\}$.
- For each $i \in \mathbb{N}$, $\varinjlim_{H \in \mathcal{U}} \prod_{n \in H} A_n \models P_i(a)$.
- For each $i \in \mathbb{N}$, $\varinjlim_{H \in \mathcal{U}} \prod_{n \in H} A_n \models \neg S_i(a)$.
- $\varinjlim_{H \in \mathcal{U}} \prod_{n \in H} A_n \models \neg Z(a)$

(It is again routine to prove that this satisfies the relevant universal property).

Note that this is different from the standard model-theoretic ultraproduct, since $Z(a)$ holds in the model-theoretic ultraproduct, but not here.

It may seem that this is a contradiction. Especially since in §7.1 it is shown that Łoś's Theorem is true in every category.

In fact, Łoś's Theorem is still true in this category **with respect to trees**. However, the axioms constraining the category mean that the trees in this category are not the same as the trees in the general category of relational structures.

In particular, there is no tree corresponding to the sentence $\forall_x Z(x)$. In the category of relational structures, such a tree could be constructed using a single object defined by:

- The underlying set is a singleton $T(\perp) = \{a\}$.
- For each $i \in \mathbb{N}$, $T(\perp) \models \neg P_i(a)$.
- For each $i \in \mathbb{N}$, $T(\perp) \models \neg S_i(a)$.
- $T(\perp) \models Z(a)$.

But this object does not satisfy the system of axioms for problematic structures.

Remark. It is necessary that $\neg P_i(a)$ and $\neg S_i(a)$ both hold in $T(\perp)$, since it is possible to have a singleton problematic structure which satisfies $Z(a)$ and $\neg P_i(a)$ for all i, and also such a structure which satisfies $Z(a)$ and $\neg S_i(a)$ for all i. Since there needs to be a morphism from $T(\perp)$ into both of these structures, then $\neg P_i(a)$ and $\neg S_i(a)$ need to hold for all i in $T(\perp)$. $\triangleleft$

So, the category-theoretic ultraproduct in the category of problematic structures still preserves all the properties that can be represented using trees in this category.

4.4.3 Remarks and Solutions

To what degree should this fact be of concern?

The following theorem shows that in many categories of interest, namely algebraic varieties, this is not a problem. This theorem is mentioned in [Lei13, Examples 8.6].

Theorem 4.4.3.1. If $\mathbb{C}$ is a category which is a variety of universal algebras, then the category-theoretic ultraproduct exists, and for any family $\langle S_x \rangle_{x \in X}$, of non-empty structures, it coincides with the model-theoretic ultraproduct.

$$\prod_{\mathcal{U}} S_x \quad \cong \quad \varinjlim_{H \in \mathcal{U}} \prod_{x \in H} S_x$$

Proof. Firstly, in a variety of universal algebras, small products and filtered colimits are calculated as in **Set**. (See, for example, [HS73, §32].) Hence also the underlying set of an ultraproduct is given as the ultraproduct of the underlying sets of the structures in the family.

It is then routine to check that the interpretations of relations, functions, and constants in the ultraproducts must also coincide. □

Corollary 4.4.3.2. In the category of relational structures $\mathbb{R}$ of some language $\mathcal{L}$, the model-theoretic ultraproduct (of non-empty structures) coincides with the category-theoretic ultraproduct.

Proof. See Lemma 1.5.0.4. □

Remark. If the standard model-theoretic definition of ultraproducts is used (where the ultraproduct is defined as the quotient of a product), then none of the structures S_x is permitted to be empty.

If, however, a definition is used where the ultraproduct is a quotient of a coproduct of a product, as mentioned in Theorem 4.3.0.2, then empty structures are permitted. ◁

Corollary 4.4.3.3. In any category that is a variety of algebras, the category-theoretic ultraproduct exists in that category, and coincides with the category-theoretic ultraproduct in the category of relational structures in the same language.

Proof. Both categories are varieties of algebras, so in both cases the category-theoretic ultraproduct exists and coincides with the model-theoretic one. Since the model-theoretic ultraproducts are isomorphic, the category-theoretic ultraproducts are also. □

The implications of this are as follows. In a general category $\mathbb{C}$ of model-theoretic structures, one needs to be careful because the ultraproduct may not quite be what is expected.

However, one can show that, in a category of algebraic structures, even though the model-theoretic and category-theoretic ultraproducts may not coincide, it is possible to extend to the category of relational structures and calculate the ultraproduct there. Łoś's Theorem implies that the ultraproduct will still be an object satisfying the correct axioms.

In fact, the above theorem implies that, if desired, it is not even necessary to extend all the way to the category of relational structures. One need simply extend to a supercategory $\mathbb{D} \supset \mathbb{C}$ of a variety of universal algebras. For example, the category of fields may be extended to the category of rings, and the ultraproduct calculated there.

4.4.4 Category-Theoretic Łoś's Theorem

The category-theoretic version of Łoś's Theorem is defined and proved in Chapter 7.

The Category-Theoretic version of Łoś's Theorem for the category of rings implies that every ultraproduct of fields calculated in the category of rings is a field. This is proved in §7.4.

However, even if the supercategory $\mathbb{D}$ is a variety, it may not contain the trees (defined in §6.2) which are required to represent the axioms for membership of $\mathbb{C}$. Hence, Łoś's Theorem in $\mathbb{D}$ might not be sufficient to prove that the ultraproduct is in $\mathbb{C}$.

Nevertheless, this is not of concern, because knowing that the ultraproduct in $\mathbb{D}$ is equivalent to the ultraproduct in the relevant category of relational structures, and knowing that Łoś's Theorem applies there, means implicitly that the ultraproduct in $\mathbb{D}$ of objects from $\mathbb{C}$ will be an object in $\mathbb{C}$.

4.4.5 Further Remarks

The paper [BN87] attempts to form a different definition of ultraproduct, in particular, one which

1. Satisfies a universal property (as opposed to the standard category-theoretic ultraproduct which is a colimit of products), in particular, the one defined is an initial object in a certain category.

2. Corresponds with the model-theoretic ultraproduct.

This paper is more successful in the first regard than the second.

Because of this, and because the proposed solutions provided above seem sufficient for our purposes, this construction is not described in more detail. However, an interested reader is encouraged to peruse it. In particular, the paper is only five pages long, and does not rely on any particularly advanced category theory or model theory.

Chapter 5

Łoś's Theorem

5.1 Łoś's Theorem - Classical

Łoś's Theorem, also known as **Łoś's Lemma** and as the **Fundamental Theorem of Ultraproducts**, is the theorem proving that a first-order sentence is true of an ultraproduct if and only if it is true for an ultrafilter-large family (a subfamily indexed by some element of the ultrafilter) of the objects composing the ultraproduct.

In fact, there exists also a version for first-order **formulae**, stated below.

Theorem 5.1.0.1 (Łoś)**.** Fix a language $\mathcal{L}$.

Let $\langle A_i \rangle_{i \in I}$ be a non-empty family of (model-theoretic) structures. Let $\mathcal{U} \subseteq \mathcal{P}(I)$ be an ultrafilter.

Let $\phi(x_1, \ldots, x_n)$ be formula of $\mathcal{L}$ (whose free variables are a subset of $\{x_1, \ldots, x_n\}$).

Then, for $a_i^j \in A_i$, the formula ϕ is true in the ultraproduct

$$\prod_{\mathcal{U}} A_i \models \phi\left(\left\langle \ \left[\left\langle a_i^1 \right\rangle_{i \in I}\right] \ , \ \ldots \ , \ \left[\langle a_i^n \rangle_{i \in I}\right] \ \right\rangle\right)$$

if and only if there is some $H \in \mathcal{U}$ such that

$$A_i \models \phi(\langle a_i^1, \ldots, a_i^n \rangle), \text{ for every } i \in H.$$

Proof. The proof is very standard, and so only a proof sketch is given. This sketch essentially follows the proof sketch given in [Ekl77]. A more complete proof may be found in [BM77, Ch 5, §3].

For any relation $R \in \mathcal{L}$, by definition of the ultraproduct (Definition 4.2.1.1), it holds that

$$R([(a_i^1)_{i \in I}], \cdots, [(a_i^n)_{i \in I}]) \qquad \text{if, and only if,} \qquad \{i \in I : R_i(a_i^1, \cdots, a_i^n)\} \in \mathcal{U}.$$

Similar arguments from definition hold for equality, constants and functions, and hence for terms and atomic formulae.

An induction argument is then performed on the logical relations $\wedge$ and $\neg$, and quantifier $\exists$.

Abbreviate

$$\bar{a} = \left\langle \ \left[\left\langle a_i^1 \right\rangle_{i\in I}\right] \ , \ \ldots \ , \ \left[\langle a_i^n \rangle_{i\in I}\right] \ \right\rangle$$

Assume the theorem is true for ψ and τ.

Then

$$\prod_{\mathcal{U}} A_i \models (\psi \wedge \tau)(\bar{a})$$

if and only if both

$$\prod_{\mathcal{U}} A_i \models \psi(\bar{a}) \qquad \text{and} \qquad \prod_{\mathcal{U}} A_i \models \tau(\bar{a})$$

which is true, by assumption of the induction, if only if there are $H_1, H_2 \in \mathcal{U}$ such that

$$A_i \models \psi(\langle a_i^1, \ldots, a_i^n \rangle), \qquad \text{for every } i \in H_1,$$

and

$$A_i \models \tau(\langle a_i^1, \ldots, a_i^n \rangle), \qquad \text{for every } i \in H_2.$$

which holds if and only if there is $H \in \mathcal{U}$ such that

$$A_i \models (\tau \wedge \psi)(\langle a_i^1, \ldots, a_i^n \rangle), \text{ for every } i \in H.$$

Similarly, assume the theorem is true for ψ, then it is shown for $\neg\psi$.

$$\prod_{\mathcal{U}} A_i \models (\neg\psi)(\bar{a})$$

if and only if

$$\prod_{\mathcal{U}} A_i \not\models \psi(\bar{a})$$

if and only if for every $H \in \mathcal{U}$ it holds that

$$A_i \not\models \psi(\langle a_i^1, \ldots, a_i^n \rangle), \text{ some } i \in H.$$

which is true if and only if

$$\{i \in I \mid A_i \models \psi(\langle a_i^1, \ldots, a_i^n \rangle)\} \notin \mathcal{U}.$$

Remark. If the above set were in $\mathcal{U}$, its intersection with any H would be in $\mathcal{U}$, but would not contain any i for which $A_i \not\models \psi(\cdots)$. $\triangleleft$

By definition of $\mathcal{U}$ being an **ultra**filter, the above is true if and only if

$$\{i \in I \mid A_i \not\models \psi(\langle a_i^1, \ldots, a_i^n \rangle)\} \in \mathcal{U}$$

and hence the above set is $H \in \mathcal{U}$ such that

$$A_i \models \neg\psi(\langle a_i^1, \ldots, a_i^n \rangle), \text{ every } i \in H.$$

(It is because of the negation that it is important to use an **ultra**filter.)

Now assume the theorem is true for $(n+1)$-ary formula $\psi(x_1, \ldots, x', \ldots, x_n)$. Then it is shown for $\exists_{x'}\psi(x_1, \ldots, x', \ldots, x_n)$.

For $a' \in \prod_{\mathcal{U}}$, let

$$a' = \left[\left\langle a_i' \right\rangle_{i \in I}\right]$$

(where $a_i' \in A_i$, for each $i \in I$) and define

$$\bar{a}' = \left\langle \; \left[\left\langle a_i^1 \right\rangle_{i \in I}\right] \; , \; \ldots \; , \; a', \ldots \; , \; \left[\langle a_i^n \rangle_{i \in I}\right] \; \right\rangle.$$

Now,

$$\prod_{\mathcal{U}} A_i \models \exists_{x'}\psi(x_1, \ldots, x', \ldots, x_n)(\bar{a})$$

if, and only if, there is some $a' \in \prod_{\mathcal{U}} A_i$ such that

$$\prod_{\mathcal{U}} A_i \models \psi(x_1, \ldots, x, \ldots, x_n)(\bar{a}')$$

if, and only if, there is some $H \in \mathcal{U}$ such that for every $i \in H$ there exists a_i' such that

$$A_i \models \psi(a_i^1, \ldots, a_i', \ldots, a_i^n)$$

which holds if, and only if, there is $H \in \mathcal{U}$ such that, for every $i \in H$, the statement

$$A_i \models \exists_{x'}\psi(a_i^1, \ldots, x', \ldots, a_i^n)$$

holds. □

Remark. The above theorem applies even in those cases where the ultrafilter $\mathcal{U}$ is principal, but the theorem becomes trivial in this case. ◁

The theorem is commonly used in the form of the following corollary:

Corollary 5.1.0.2. If ϕ is a first-order **sentence**, then

$$\prod_{\mathcal{U}} A_i \models \phi$$

if and only if there is some $H \in \mathcal{U}$ such that

$$A_i \models \phi, \text{ for every } i \in H.$$

◁

A direct consequence of this is the following corollary:

Corollary 5.1.0.3. Let $\mathcal{T}$ be a first-order theory, and let $\langle A_i \rangle_{i \in I}$ be a family of structures such that

$$A_i \models \mathcal{T} \qquad \text{for every } i \in I.$$

Then, for any ultrafilter $\mathcal{U} \subseteq \mathcal{P}(I)$, the ultraproduct

$$\prod_{\mathcal{U}} A_i \models \mathcal{T}.$$

◁

In particular, this means that an ultraproducts of groups is a group, an ultraproduct of rings is a ring, an ultraproduct of lattices is a lattice, and so forth.

As noted in [Ekl77], this is a useful way to prove a certain theory is not first-order axiomatizable. In §2.2.4 of this dissertation above, it is shown, for example, that the property of a field having characteristic 0 is not first-order axiomatizable.

5.2 Definable Subsets of Ultraproduct

5.2.1 Overview

The approach in this section was brought to my attention by my supervisor, Dr Gareth Boxall, though they may be well-known in the literature.

The idea is as follows. Given a family $\langle A_i \rangle_{i \in I}$ of structures, consider the definable subsets.

For example, restricting our attention to an n-ary relation R, the definable subset of A_i corresponding to R is the subset of $(A_i)^n$ of tuples which satisfy R (in strict set-theoretic terms, the relation R is usually defined to actually be exactly this subset).

From the definable subsets corresponding to the interpretations of symbols in the language, the definable subset corresponding to a formula may be built up using the set-theoretic operations of union, intersection, complement and projection. These corresponding to the logical operations of 'or', 'and', 'not' and 'exists' respectively.

Given an ultraproduct of the family, definable subsets of the interpretations of symbols in the language correspond, by definition, to their definition in 'large subfamilies'. For example, an n-tuple of equivalence classes is in the definable subset of the ultraproduct corresponding to the n-ary relation symbol R if, and only if, there is an ultrafilter-large subfamily of $\langle A_i \rangle_{i \in I}$ such that the n-tuples, with components coming from the equivalence classes, are in the definable subsets of R in each structure in this subfamily.

If the set-theoretic operations preserve this correspondence, then also the definable subsets for all formulae will correspond in the same way. In essence, this is Łoś's Theorem.

5.2.2 Definitions and Theorem

Let $\langle A_i \rangle_{i \in I}$ be a family of sets indexed by I. Fix $n \in \mathbb{N}$ and for each $i \in I$ let $X_i \subseteq \langle A_i \rangle^n$.

Let $\mathcal{U}$ be an ultrafilter on $\mathcal{P}(A)$ and consider the subset of the n-th power of the ultraproduct

$$[\langle X_i \rangle_{i \in I}] \subseteq \left(\prod_{\mathcal{U}} A_i \right)^n$$

defined as

Definition 5.2.2.1.

$$[\langle X_i \rangle_{i \in I}] =$$
$$\left\{ \langle [\langle x_i^1 \rangle_{i \in I}], \ldots, [\langle x_i^n \rangle_{i \in I}] \rangle \in \left(\prod_{\mathcal{U}} A_i \right)^n \middle| \{ i \in I \mid \langle x_i^1, \ldots, x_i^n \rangle \in X_i \} \in \mathcal{U} \right\}$$

$\lhd$

Remark. The reader is welcome to check for themselves that the above definition is well-defined, in the sense that the question of membership of a given tuple of equivalence classes is not dependent on the specific representatives chosen. $\lhd$

Now let $\phi(\bar{x})$ be a formula in (at most) n variables and let each X_i be the subset of A_i defined by $\phi(\bar{x})$:

$$X_i = \{ \langle x_i^1, \ldots, x_i^n \rangle \in A_i^n \ : \ \phi(x_i^1, \ldots, x_i^n) \}$$

Then

Theorem 5.2.2.2. $[\langle X_i \rangle_{i \in I}]$ is the subset of $\prod_{\mathcal{U}} A_i$ defined by $\phi(\bar{x})$. $\lhd$

To show this means to show that

$$\left\langle [\langle x_i^1 \rangle_{i \in I}], \ldots, [\langle x_i^n \rangle_{i \in I}] \right\rangle \in [\langle X_i \rangle_{i \in I}]$$

which by definition is equivalent to

$$\{i \in I \ : \ \langle x_i^1, \dots, x_i^n \rangle \in X_i\} \in \mathcal{U}$$

if, and only if,

$$\phi\Big([\langle x_i^1 \rangle_{i \in I}], \dots, [\langle x_i^n \rangle_{i \in I}]\Big).$$

This is just Łoś's Theorem in a different form. It states that the formula ϕ holds in the ultraproduct for the tuple $\langle [\langle x_i^1 \rangle_{i \in I}], \dots, [\langle x_i^n \rangle_{i \in I}] \rangle$ of equivalence classes, if, and only if, the set of $i \in I$ for which ϕ holds for the tuples $\langle x_i^1, \dots, x_i^n \rangle$ is in the ultrafilter.

Lemma 5.2.2.3. If X_i is the interpretation of the n-ary relation R then $[\langle X_i \rangle_{i \in I}]$ is the interpretation of R in the ultraproduct.

Proof. By definition of how relations are interpreted in the ultraproduct.

$$\Big\langle [\langle x_i^1 \rangle_{i \in I}], \cdots, [\langle x_i^n \rangle_{i \in I}] \Big\rangle \in [\langle X_i \rangle_{i \in I}],$$

if, and only if,

$$\{i \in I : (x_i^1, \cdots, x_i^n) \in X_i\} \in \mathcal{U},$$

if, and only if,

$$\{i \in I : R(x_i^1, \cdots, x_i^n)\} \in \mathcal{U},$$

if, and only if,

$$R\Big([\langle x_i^1 \rangle_{i \in I}], \cdots, [\langle x_i^n \rangle_{i \in I}]\Big).$$

□

The current section will only concern relations. The cases for functions and constants can be shown in a similar manner, or can be shown to follow from the case for relations via a translation of theories, as in Definition 1.5.0.2.

Similarly, equality may also be viewed as a special case of the above, namely via the binary relation corresponding to equality.

Given the above lemma, Łoś's theorem hence reduces to the preservation in the ultraproduct of intersections, unions, complements and projections as follows.

Lemma 5.2.2.4.

$$[\langle X_i \rangle_{i \in I}] \cap [\langle Y_i \rangle_{i \in I}] = [(X_i \cap Y_i)_{i \in I}]$$

$$[\langle X_i \rangle_{i \in I}] \cup [\langle Y_i \rangle_{i \in I}] = [(X_i \cup Y_i)_{i \in I}]$$

$$\left(\prod_{\mathcal{U}} A_i\right)^n - [\langle X_i\rangle_{i\in I}] = [(M_a^n - X_i)_{i\in I}]$$

$$\pi_{\not{i}}([\langle X_i\rangle_{i\in I}]) = \left[(\pi_{\not{i}}\langle X_i\rangle)_{i\in I}\right],$$

where $\pi_{\not{i}}$ is the projection map forgetting coordinate i:

$$\pi_{\not{i}}(x_1, \ldots, x_n) = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n).$$

Proof. The following is merely a proof sketch.

If follows directly from the properties of ultrafilters that,

$$\{i \in I \mid \langle x_i^1, \ldots, x_i^n\rangle \in X_i \cap Y_i\} \in \mathcal{U}$$

if, and only if, both

$$\{i \in I \mid \langle x_i^1, \ldots, x_i^n\rangle \in X_i\} \in \mathcal{U} \quad \text{and} \quad \{i \in I \mid \langle x_i^1, \ldots, x_i^n\rangle \in Y_i\} \in \mathcal{U}.$$

Thus

$$\begin{aligned}
&[\langle X_i\rangle_{i\in I}] \cap [\langle Y_i\rangle_{i\in I}] \\
&= \left\{ \langle [\langle x_i^1\rangle_{i\in I}], \ldots, [\langle x_i^n\rangle_{i\in I}]\rangle \in \left(\prod_{\mathcal{U}} A_i\right)^n \right. \\
&\qquad\qquad \left. \middle| \; \{i \in I \mid \langle x_i^1, \ldots, x_i^n\rangle \in X_i \cap Y_i\} \in \mathcal{U} \right\} \\
&= [(X_i \cap Y_i)_{i\in I}]
\end{aligned}$$

Similar arguments work for the rest of the preservation properties. □

These set operations correspond to the logical operations **and**, **or**, **not** and **there exists** respectively, from which all formulae of first order logic can be built. This provides the necessary prerequisites to show that the (definable) subset of the ultraproduct defined by a given formula ϕ is an "equivalence class" $[\langle X_i\rangle_{i\in I}]$ of the family $\langle X_i\rangle_{i\in I}$ of the corresponding (definable) subsets of the $\langle A_i\rangle_{i\in I}$ defined by the same formula ϕ.

Łoś's Theorem then follows as an easy corollary.

Proof of Theorem 5.2.2.2. The proof goes by induction. The example for **and** is given. The other operations follow similarly.

For each formula ϕ, define the set

$$X_\phi = \{\langle x_i^1, \ldots, x_i^n\rangle \in A_i^n \ : \ \phi(x_i^1, \ldots, x_i^n)\}$$

Assume the theorem holds for ψ and τ, both of whose free variables are a subset of $\{x_1, \ldots, x_n\}$.

Then, for each $i \in I$,

$$X_\psi = \{\langle x_i^1, \ldots, x_i^n\rangle \in A_i^n \ : \ \psi(x_i^1, \ldots, x_i^n)\},$$

and

$$X_\tau = \{\langle x_i^1, \ldots, x_i^n\rangle \in A_i^n \ : \ \tau(x_i^1, \ldots, x_i^n)\},$$

so

$$X_\psi \cap X_\tau = X_{\psi\wedge\tau} = \{\langle x_i^1, \ldots, x_i^n\rangle \in A_i^n \ : \ \psi(x_i^1, \ldots, x_i^n) \wedge \tau(x_i^1, \ldots, x_i^n)\}.$$

Thus, by preservation of such sets in the ultraproducts,

$$[(X^i_{\psi\wedge\tau})_{i\in I}] = [(X^i_\psi \cap X^i_\tau)_{i\in I}] = [(X^i_\psi)_{i\in I}] \cap [(X^i_\tau)_{i\in I}].$$

Similar properties hold for the other operations. □

Corollary 5.2.2.5. See Theorem 5.1.0.1.

Proof.

$$\prod_{\mathcal{U}} I_i \models \phi\left(\left\langle \ \left[\left\langle a_i^1\right\rangle_{i\in I}\right] , \ \ldots \ , \ \left[\langle a_i^n\rangle_{i\in I}\right] \ \right\rangle\right)$$

if, and only if,

$$\left\langle\left[\left\langle a_i^1\right\rangle_{i\in I}\right] , \ \ldots \ , \ \left[\langle a_i^n\rangle_{i\in I}\right]\right\rangle \in \left[\left(X^i_\phi\right)_{i\in I}\right],$$

if, and only if,

$$\{i \in I \mid \langle x_i^1, \ldots, x_i^n\rangle \in X_i\} \in \mathcal{U},$$

if, and only if, there is some $H \in \mathcal{U}$ such that

$$A_i \models \phi(\langle a_i^1, \ldots, a_i^n\rangle), \text{ for every } i \in H.$$

□

Chapter 6

Trees to Represent First-Order Formula

6.1 Introduction

6.1.1 Overview

This chapter and the next follow the results of the papers [AN78] and [AN79]. The paper [AN79] defines a category-theoretic means of representing first-order formulae, and [AN78] provides an entirely category-theoretic proof that category-theoretic ultraproducts (as defined in §4.3), satisfy Łoś's Theorem with respect to this category-theoretic version of formulae.

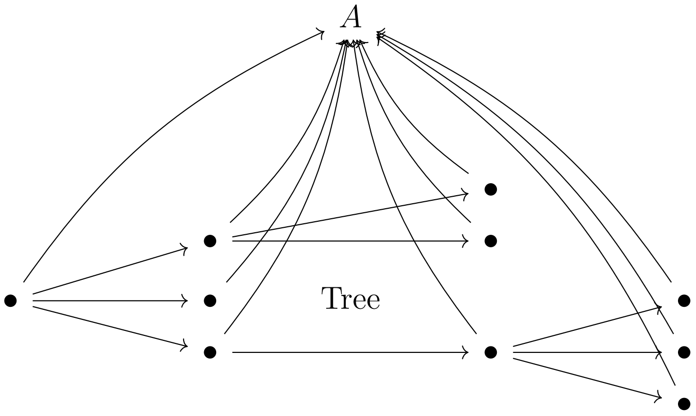
Remark. The author of this dissertation is uncertain why the former paper is published after the latter. However, all of the concepts necessary for the proof of Łoś's Theorem are defined in [AN78], and discussed in more detail in [AN79]. $\triangleleft$

In the current chapter, the definitions and theorems of [AN79] are described in detail.

Notably, the definitions, theorems, and proofs of §6.4 are due to the current author. In [AN79], it is only stated that it is possible to construct a tree from a given formula, and the properties of such a tree are described. However, the exact construction for trees from formulae appears (to the author's knowledge) for the first time in this dissertation.

6.1.2 Basic Concepts

Given a signature L consisting of relations (and equality), each sentence ψ of first-order logic can be encoded as a tree of structures under this signature. Given another such structure A, the truth of ψ in A is equivalent to the existence of morphisms from the tree into A satisfying certain properties.



Fix a signature L consisting only of finitary relational symbols. The category of relational structures of L is a category whose objects are sets equipped with an n-ary relation for each n-ary relational symbol, and whose morphisms are maps which preserve relations (in other words, if $f : X \to Y$ is a morphism, then if $x_1, \ldots, x_n$ are R-related in X, then $f(x_1), \ldots, f(x_n)$ are R-related in Y).

Intuitively, a tree is a collection of relational structures and morphisms which forms a tree shape. Some of the objects and/or morphisms may be repeated.

$$\begin{tikzcd}
 & & X_2 & \\
 & X_2 \arrow[ur, "1_{X_2}"] \arrow[r, "m_2"] & X_3 & \\
X_1 \arrow[ur, "m_1"] \arrow[r, "m_1"] \arrow[dr, "m_3"'] & X_2 & & X_1 \\
 & X_4 \arrow[r, "m_4"] & X_5 \arrow[ur, "m_5"] \arrow[r, "m_6"] \arrow[dr, "m_7"'] & X_6 \\
 & & & X_2
\end{tikzcd}$$

Formally, such a tree is given as a functor from a category shaped like a tree (a poset which is a rooted tree) into the category of relational structures.

Given a relational structure A, the sentence ψ will be true of A if and only if A satisfies a certain condition with respect to the tree encoding ψ, called **injectivity**.

The object A is injective with respect to the tree T if for every morphism $f_1 : X_1 \to A$, there exists some triple (X_2, m_2, e_2) (one level beyond it) such that $e_2 \circ m_2 = f_1$ and for every triple (X_3, m_3, f_3) (one level beyond that) where $f_3 \circ m_3 = e_2$, there exists some (X_4, m_4, e_4) (one level beyond that) such that $e_4 \circ m_4 = f_3$ etc. (for a finite number of iterations) up until the end of some branch of the tree.

$$\begin{tikzcd}
& & A & & \\
X_1 \arrow[urr, bend left, "\forall f_1"] \arrow[r, "\exists m_2"] & \exists X_2 \arrow[ur, "\exists e_2"] \arrow[r, "\forall m_3"] & \forall X_3 \arrow[u, "\forall f_3"'] \arrow[r, "\exists m_4"] & \exists X_4 \arrow[ul, "\exists e_4"'] \arrow[r] & \cdots \arrow[ull, bend right]
\end{tikzcd}$$

Conversely, given a tree T, it is also possible to find a sentence ψ so that, again, ψ is true of A if and only if A is injective with respect to T.

If a sentence ψ is encoded as a tree T, and tree is converted back, the sentence ψ' that arises may not be equal to ψ. However, they will be logically equivalent.

Two trees may be considered equivalent if they are injective with respect to all of the same objects. Then, again, if a tree T is converted to a sentence ψ and then converted back, the tree T' may not be equal to T, but it will be equivalent in this sense.

Due to Gödel's theorem of the completeness of first-order logic, it then follows that two sentences ψ and ψ' are equivalent if and only if their resulting trees T and T' are equivalent, and similarly that two trees T and T' are equivalent if and only if their resulting sentences ψ and ψ' are equivalent.

6.1.3 Example of Tree and Corresponding Formula

Assume we are working in a relational language with two relations R and S. Then the statement

$$\forall_{x,y}(R(x,y) \Rightarrow S(x,y))$$

corresponds to a tree

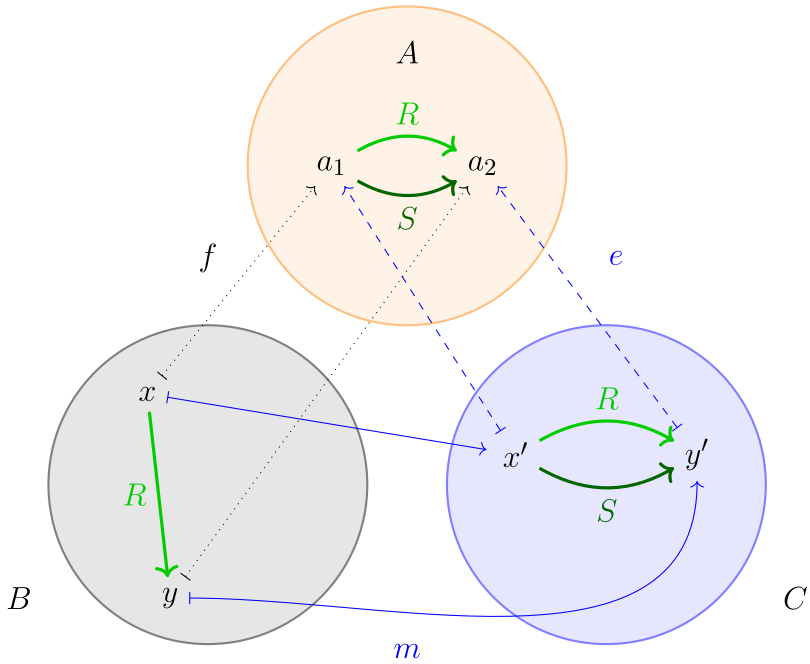
$$B \xrightarrow{m} C$$

where $B = \{x, y\}$ and $\langle x, y\rangle \in R$ but nothing else is related and $C = \{x', y'\}$ with $\langle x', y'\rangle \in R$ and $\langle x', y'\rangle \in S$ but nothing else is related. The morphism m maps $x \mapsto x'$ and $y \mapsto y'$.

Injectivity of a structure A with respect to this tree means that for every **morphism** $f : B \to A$ there exists a **morphism** $e : C \to A$ such that $em = f$.

$$
\begin{tikzcd}
 & A & \\
B \arrow[ur, "f"] \arrow[rr, "m"] & & C \arrow[ul, "e"']
\end{tikzcd}
$$

This may be clearer in a picture. Here an arrow $R : x \to y$, for example, indicates that x is R-related to y.



Now, the map f corresponds to choosing two elements of A. Furthermore, the fact that it is a morphism means that a_1 and a_2 are R-related. Injectivity of A with respect to the tree means existence of an e for every f.

Now, the map e corresponds also to choosing two elements of A. Again, the fact that it is a morphism means that these two elements must be R and S-related. Injectivity implies the existence of such an e (for which $f = em$). Hence it implies $\exists_{x',y'} R(x', y') \wedge S(x', y')$.

However, the fact that $em = f$ means that $x = x'$ and $y = y'$. Hence the formula is

$$\forall_{x,y}[R(x, y) \Rightarrow \exists_{x',y'}(x = x' \wedge y = y' \wedge R(x', y') \wedge S(x', y'))]$$

which is equivalent to

$$\forall_{x,y}[R(x, y) \Rightarrow R(x, y) \wedge S(x, y)]$$

which is itself equivalent to

$$\forall_{x,y}[R(x,y) \Rightarrow S(x,y)].$$

Hence any A which is injective with respect to the tree must satisfy the formula.

Note in particular that if A is empty, then there do not exist any $f : B \to A$, so the "for every morphism $f \ldots$" aspect of the injectivity is trivially satisfied. Similarly, $\forall_{x,y} \ldots$ is trivially satisfied. So the definition still works.

6.2 Trees

6.2.1 Tree Categories

Definition 6.2.1.1. We call a category $\mathbb{T}$ a **tree category** if it is a finite poset which is a (rooted) tree in which all meets exist, and for any pair of objects x and y, if they share an upper bound $z \geq x, y$, then x and y must already be comparable ($x \leq y$ or $y \leq x$).

In other words, as a category:

- There are only finitely many objects.
- There is a unique (not just up to isomorphism) initial object $\perp = \perp_T \in \mathbb{T}$.
- For any two objects $X, Y \in \mathbb{T}$, there is either no morphism or a unique morphism (called 'less or equal') $x_Y : X \to Y$.
- For any $X, Y \in \mathbb{T}$, if there are morphisms $x_Y : X \to Y$ and $y_X : Y \to X$ then $X = Y$ and $x_Y = y_x = 1_X = 1_Y$.
- For any $X, Y \in \mathbb{T}$, there is object $Z \in \mathbb{T}$ ('meet of X and Y'), and morphisms $z_X : Z \to X$ and $z_Y : Z \to Y$, which is universal in the sense that given any other $W \in \mathbb{T}$, and morphisms $w_X : W \to X$ and $w_Y : W \to Y$, there is $w_Z : W \to Z$ with $z_X \circ w_Z = w_X$ and $z_Y \circ w_Y = w_Y$.
- For any pair of objects $X, Y \in \mathbb{T}$, if there is Z and morphisms $x_Z : X \to Z$ and $y_Z : Y \to Z$, then there is already either $x_Y : X \to Y$ or $y_X : Y \to X$.

The initial object is also called the **root** of the tree-category. $\triangleleft$

The diagram below on the left shows an example of a category which is a tree (with morphisms in different styles for clarity). However, usually when a tree is drawn, many of the morphisms are left out, so that it forms an actual graph-theoretic tree, with the compositions forming the undrawn morphisms implicit. In both of the diagrams, the identity morphisms are also left undrawn.

$$
\begin{tikzcd}
 & & Y_1^1 & \\
 & X_1 \arrow[ur] \arrow[r] & Y_2^1 & \\
\perp \arrow[ur] \arrow[r] \arrow[dr] \arrow[uurr, dotted, bend left] \arrow[urr, dotted] \arrow[rrr, dotted] \arrow[drr, dotted] \arrow[drrr, dotted] \arrow[ddrrr, dotted, bend right] & X_2 & & Z_1^{3,1} \\
 & X_3 \arrow[r] \arrow[urr, dashed] \arrow[rr, dashed, bend right] \arrow[drr, dashed, bend right] & Y_1^3 \arrow[ur] \arrow[r] \arrow[dr] & Z_2^{3,1} \\
 & & & Z_3^{3,1}
\end{tikzcd}
\qquad
\begin{tikzcd}
 & & Y_1^1 & \\
 & X_1 \arrow[ur] \arrow[r] & Y_2^1 & \\
\perp \arrow[ur] \arrow[r] \arrow[dr] & X_2 & & Z_1^{3,1} \\
 & X_3 \arrow[r] & Y_1^3 \arrow[ur] \arrow[r] \arrow[dr] & Z_2^{3,1} \\
 & & & Z_3^{3,1}
\end{tikzcd}
$$

Definition 6.2.1.2. Given a tree category $\mathbb{T}$, recursively define the **layer** of an object $X \in \mathbb{T}$ as follows:

- $\text{Layer}(\perp) = 1$.
- $\text{Layer}(X) = \max\{\text{Layer}(Y) : Y \in \mathbb{T}\ ,\ Y \neq X\ ,\ \exists f : Y \to X\} + 1$

An object X with $\text{Layer}(X) = n$ is said to be an **object at layer** n, or a **layer** n **object**. ◁

Remark. An object's layer in the tree category is 'how many steps it is from the root' (initial object), plus one. ◁

Definition 6.2.1.3. Given tree category $\mathbb{T}$ and $X \in \mathbb{T}$, the **sub-tree-category** $\mathbb{S}_{\mathbb{T}}(X) \subseteq \mathbb{T}$ **with root** X is a subcategory which is itself a tree category with root $\perp_S = X$ defined as follows.

It is the full subcategory containing all codomains of morphisms from X.

In more detail, it contains X, all morphisms $x_Y : X \to Y$ along with their codomains Y, all morphisms $y_Z : Y \to Z$ along with their codomains Z and so forth.

It is equipped with category inclusion $I : \mathbb{S}_{\mathbb{T}}(X) \hookrightarrow \mathbb{T}$. ◁

The diagram below shows the subtree of the given tree with root X_3.

$$
\begin{tikzcd}
 & & Y_1^1 & \\
 & X_1 \arrow[ur, dotted] \arrow[r, dotted] & Y_2^1 & \\
\perp \arrow[ur, dotted] \arrow[r, dotted] \arrow[dr, dotted] & X_2 & & Z_1^{3,1} \\
 & X_3 \arrow[r] & Y_1^3 \arrow[ur] \arrow[r] \arrow[dr] & Z_2^{3,1} \\
 & & & Z_3^{3,1}
\end{tikzcd}
\qquad
\begin{tikzcd}
 & & Z_1^{3,1} \\
X_3 \arrow[r] & Y_1^3 \arrow[ur] \arrow[r] \arrow[dr] & Z_2^{3,1} \\
 & & Z_3^{3,1}
\end{tikzcd}
$$

6.2.2 Tree Functors

Definition 6.2.2.1. Given a category $\mathbb{A}$, a **tree** T in $\mathbb{A}$ is a functor $T : \mathbb{T} \to \mathbb{A}$ where $\mathbb{T}$ is a tree category. $\triangleleft$

Definition 6.2.2.2. Given tree $T : \mathbb{T} \to \mathbb{A}$ and $X \in \mathbb{T}$, the **sub-tree** $S_T(X) \subseteq T$ **with root** X is the functor $S_T(X) : \mathbb{S}_{\mathbb{T}}(X) \to \mathbb{A}$ such that $S_T(X) = TI$ and $S_T(X)$ has root $\perp_S = X$. $\triangleleft$

Remark. We are mainly interested in the objects in the image of such a tree. However, we cannot refer directly to the image, because we may want to use many copies of the same object (and morphism) in different portions of the tree. Hence, the above definition is the right one. $\triangleleft$

Figure 6.1 shows an example of a tree in the (poset) category of natural numbers. Note how some objects may be repeated in the image of the tree.

6.3 Injectivity of a Tree

6.3.1 Motivation of Definition

This is an overview of the definition of injectivity of a tree. The formal definitions will be given in the following subsections.

An object A is injective with respect to a tree T if for every morphism $f_1 : T(\perp) \to A$, there exists some triple (X_2, m_2, e_2) where $e_2 \circ T(m_2) = f_1$ and for every triple (X_3, m_3, f_3) where $f_3 \circ T(m_3) = e_2$, there exists some (X_4, m_4, e_4) where $e_4 \circ T(m_4) = f_3$ etc. (for a finite number of iterations) up until the end of some branch of the tree.

$$
\begin{tikzcd}
 & & Y_1^1 & \\
 & X_1 \arrow[ur] \arrow[r] & Y_2^1 & \\
\perp \arrow[ur] \arrow[r] \arrow[dr] & X_2 & & Z_1^{3,1} \\
 & X_3 \arrow[r] & Y_1^3 \arrow[ur] \arrow[r] \arrow[dr] & Z_2^{3,1} \\
 & & & Z_3^{3,1}
\end{tikzcd}
\qquad
\begin{tikzcd}
 & & 2 & \\
 & 2 \arrow[ur] \arrow[r] & 3 & \\
1 \arrow[ur] \arrow[r] \arrow[dr] & 2 & & 8 \\
 & 5 \arrow[r] & 6 \arrow[ur] \arrow[r] \arrow[dr] & 8 \\
 & & & 9
\end{tikzcd}
$$

Figure 6.1: Tree in Category of Natural Numbers

Remark. In the diagram below, the use of quantifiers is abused in order to indicate whether the definition requires a property for at least one such morphism or for all such morphisms. ◁

$$
\begin{tikzcd}
 & & A & & \\
T(\perp) \arrow[urr, bend left, "\forall f_1"] \arrow[r, "T(\exists m_2)"] & T(\exists X_2) \arrow[ur, "\exists e_2"] \arrow[r, "T(\forall m_3)"] & T(\forall X_3) \arrow[u, "\forall f_3"'] \arrow[r, "T(\exists m_4)"] & T(\exists X_4) \arrow[ul, "\exists e_4"'] \arrow[r] & \cdots \arrow[ull, bend right]
\end{tikzcd}
$$

Remark. The notation of f and e for the morphisms has been chosen to correspond to 'for all' and 'exists' respectively. The subscript i in f_i or e_i corresponds to the subscript of X_i occurring in the domain, whereas in m_i it corresponds to the subscript of X_i occurring in the codomain. I.e., for any triple, (X_i, m_i, f_i) or (X_i, m_i, e_i) the subscripts are the same for each component of the triple. ◁

In the category of relational structures, this alternation of 'for all' and 'exists' of morphisms will allow the representation of certain first-order formulae beginning with alternating $\forall\exists$, i.e., of the form

$$\forall_{x_1}\exists_{x_2}\forall_{x_3}\exists_{x_4}\ldots\forall_{x_{2m-1}}\exists_{x_{2m}}(\phi(x_1,\ldots x_{2m})),$$

where ϕ is a propositional formula.

This can then be made to represent a formula starting with any finite string of $\forall$s and $\exists$s in any order (including repetitions) by not mentioning some of the variables x_i in the formula ϕ. For example,

$$\forall_{x_1}\exists_{x_2}\forall_{x_3}\exists_{x_4}(x_1 = x_3) \qquad \Leftrightarrow \qquad \forall_{x_1}\forall_{x_3}(x_1 = x_3).$$

Remark. The above is not strictly true since there are exceptions with empty structures. Note that $\forall_{x_1}\exists_{x_2}(x_2 = x_2)$ is true for an empty model whereas $\exists_{x_2}(x_2 = x_2)$ is false for such a model. This does not affect §6.4, and for §6.5 some explanation is given in this regard in §6.5.6. $\lhd$

Furthermore, it will be shown there is a way to choose structures in the image of the tree so that **any** ϕ can be represented, provided it is of the form

$$\bigwedge_{i<z}(\alpha_i \rightarrow \bigvee_{j<p_i} \beta^i_j)$$

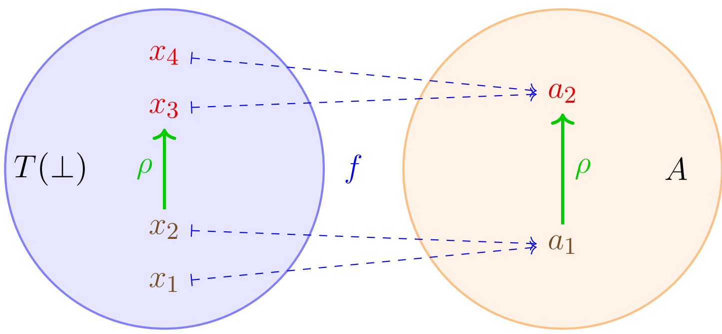
where the α_i are conjunctions of atomic formulae and β^i_j are atomic formulae. Then, by [HMT71], this is equivalent to any propositional formula, and so any formula can be represented.

6.3.2 Intuition for Definition

The object $T(\perp)$ should be interpreted as a collection of variables. A morphism $f_1 : T(\perp) \rightarrow A$ is an assignment of these variables to elements of A.

Then injectivity of the object A with respect to the tree T under the assignment f is related to truth of the formula ϕ under the assignment of the free variables in ϕ by $f(x_1), \ldots, f(x_n)$ in A.

However, there may be relations holding between some elements of $T(\perp)$. Then each morphism $f : T(\perp) \rightarrow A$ actually corresponds to a assignment of variables **such that** the relevant relations hold between the assignments. See, for example, the picture below which details a specific assignment $f : T(\perp) \rightarrow A$. Then the fact that x_2 and x_3 are related by ρ means that a_1 and a_2 must be related for f to be morphism.



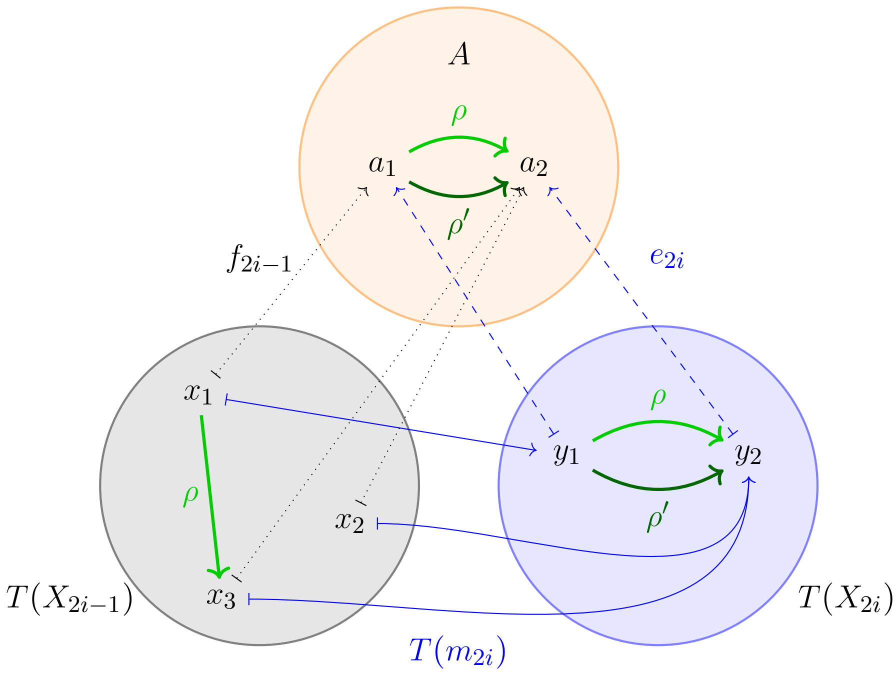
Each morphism from the other objects into A is also an assignment of variables (with specific relations).

Each object can introduce new variables and declare previous variables equal, as well as introduce new relations.

Note that all of the relations between elements of some object of the tree must also hold between the images of those elements under some $T(m)$.

For example, in the diagram below, since x_1 is related by ρ to x_3, then for $T(m_{2i})$ to be a morphism means that ρ must relate y_1 and y_2. Similarly, for e_{2i} to be a morphism means that both ρ and ρ' must relate a_1 and a_2.

For the morphisms to commute, the images of x_2 and x_3 must be equal in A since their images are equal (to y_2) in $T(X_{2i})$.



Now, injectivity of A with respect to T does not imply all of the above being true. Instead, injectivity in this case means that for every assignment x_1, x_2 and x_3 in A such that x_1 is ρ-related to x_3, there is some assignment y_1, y_2 in A such that $x_1 = y_1$ and $x_2 = x_3 = y_2$ and y_1 is both ρ and ρ' related to y_2.

6.3.3 Definitions

Definition 6.3.3.1. Given a category $\mathbb{A}$ and a tree $T : \mathbb{T} \to \mathbb{A}$, an object $A \in \mathbb{A}$ and morphism $f : T(\perp) \to A$.

Then A is **injective** with respect to T, under the evaluation f denoted $A \models T[f]$, if it satisfies the following recursive condition:

There exist

- Object $X_2 \in \mathbb{T}$ with $\mathrm{Layer}(X_2) = 2$,

- Morphism $m : \perp \to X_2$ in $\mathbb{T}$,
- Morphism $e : T(X_2) \to A$ in $\mathbb{A}$ with $e \circ T(m) = f$

such that for any

- Object $X_3 \in \mathbb{T}$ with $\text{Layer}(X_3) = 3$,
- Morphism $k : X_2 \to X_3$ in $\mathbb{T}$,
- Morphism $d : T(X_3) \to A$ in $\mathbb{A}$ with $d \circ T(k) = e$

it holds that

$$A \models S_T(X_3)[d]$$

where $S_T(X_3)$ is the subtree of T with root X_3.

$$\begin{tikzcd}
 & A & \\
T(\perp) \arrow[ur, "f"] \arrow[r, "T(m)"'] & T(X_2) \arrow[u, "e"'] \arrow[r, "T(k)"'] & T(X_3) \arrow[ul, "d"']
\end{tikzcd}$$

◁

Remark. Note in particular if there are object X_2, and morphisms $m : \perp \to X_1$ and $e : T(X_2) \to A$ with $e \circ T(m) = f$ but for which there is no object X_3, and morphisms $k : X_2 \to X_3$ and $d : T(X_3) \to A$ with $d \circ T(k) = e$, then the condition that $A \models S_T(X_3)[d]$ for all suitable d is trivially satisfied (since it is satisfied for all of the non-existent d), and so it would hold that $A \models T[f]$. ◁

Definition 6.3.3.2. The object A is **injective** with respect to T, denoted $A \models T$, if $A \models T[f]$ for all morphisms $f : T(\perp) \to A$. ◁

Remark. Again, if there is no morphism $f : T(\perp) \to A$, then the above condition is trivially satisfied. ◁

6.4 First-Order Formulae from Trees of Finite Relational Structures

6.4.1 Notation

In this section, the following notation is used for different layers of the tree Here, a variable with a hat, (like $\hat{t}_1$), indicates the maximum value that the variable (t_1) can take.

$$
\begin{tikzcd}
 & & X_3^{1,1} & \\
 & & \vdots & \\
 & X_2^1 \arrow[uur] \arrow[ur] \arrow[r] & X_3^{1,\hat{t}_1} & X_3^{q,1} \\
 & \vdots & & \vdots \\
\perp = X_1 \arrow[uur] \arrow[ur] \arrow[r] \arrow[dr] \arrow[ddr] & X_2^q \arrow[uurr] \arrow[urr] \arrow[rr] \arrow[drr] \arrow[ddrr] & & X_3^{q,t_q} \\
 & \vdots & & \vdots \\
 & X_2^{\hat{q}} \arrow[r] \arrow[dr] \arrow[ddr] & X_3^{\hat{q},1} & X_3^{q,\hat{t}_q} \\
 & & \vdots & \\
 & & X_3^{\hat{q},\hat{t}_{\hat{q}}} &
\end{tikzcd}
$$

A simple specific example is the following

$$
\begin{tikzcd}
 & & X_3^{1,1} \\
 & X_2^1 \arrow[ur] \arrow[r] & X_3^{1,2} \\
\perp \arrow[ur] \arrow[r] \arrow[dr] & X_2^2 & \\
 & X_2^3 \arrow[r] & X_3^{3,1}
\end{tikzcd}
$$

Remark. Note that the tree itself may extend beyond the third layer, but for the purposes of this section, the notation up to the third layer will be sufficient. Reference to further layers is captured by the subtree notation $S_T(X_3^{q,t_q})$. $\lhd$

6.4.2 Intuition of Formula Given Tree

Although the definitions of the previous section are valid even when using infinite relational structures (those which have an underlying set which is infinite in size), first-order formulae correspond exactly to the trees consisting of finite relational structures. The object A, however, may still be infinite.

Fix a finite signature Σ and let **RStr** be the category of relational structures in this signature (in this section referred to as 'structures'). Then, given a tree $T : \mathbb{T} \to \mathbf{RStr}$, whose image consists of **finite** relational structures, a formula $\phi_T(x_1, \ldots, x_d)$ based on T can be defined so that

$$\forall_{x_1,\ldots,x_d \in A}[\phi_T(x_1, \ldots, x_d)] \qquad \text{iff} \qquad A \models T.$$

Remark. The above ϕ_T may be a sentence (a formula in no free variables), in which case the $\forall_{x_1,\ldots,x_d}$ has no variables and so is simply left off. $\lhd$

The motivation for the definition given below can be found by looking at the proof of Theorem 6.4.4.1. It is recommended that the definition and theorem be read in parallel. However, in this subsection, a brief description of each component of the definition is given.

Remark. In this section, including the proofs, everything is stated as though A and $T(X)$ for each X were non-empty. Everything still works for empty structures, in most places as a trivial special case. Nevertheless, linguistically it is easier to speak of elements, even though they may not exist. In some places remarks are used to indicate the type of changes necessary to extend the definitions and proof for empty structures. $\lhd$

Given the definition of the formula $\phi_{S_T(X_3^{q,t_q})}\left(z_1, \ldots, z_{d_{t_q}}\right)$ for each subtree $S_T(X_3^{q,t_q})$, the formula $\phi_T(x_1, \ldots, x_d)$ is defined recursively by:

$$\bigwedge_{\rho,\bar{x}} \alpha_{\rho,\bar{x}} \Rightarrow \bigvee_q \left[\exists_{y_1,\ldots,y_{n_q}} \left(\bigwedge_{\rho,\bar{y}} \beta^q_{\rho,\bar{y}} \wedge \bigwedge_{r_q} \left(x_{i_{r_q}} = y_{j_{r_q}}\right) \wedge \bigwedge_{t_q} \delta_{t_q}\right)\right]$$

where $\delta_{t_q} \equiv$

$$\forall_{z_1,\ldots,z_{d_{t_q}}} \left[\left(\bigwedge_{s_{t_q}} \left(y_{i_{s_{t_q}}} = z_{j_{s_{t_q}}}\right) \wedge \bigwedge_{\rho,\bar{z}} \gamma^{q,t_q}_{\rho,\bar{z}}\right) \quad \Rightarrow \quad \phi_{S_T(X_3^{q,t_q})}\left(z_1, \ldots, z_{d_{t_q}}\right)\right].$$

Remark. The numbers n_q and d_{t_q} may be zero, in which case the existential quantifier $\exists_{y_1,\ldots,y_{n_q}}$ and universal quantifier $\forall_{z_1,\ldots,z_{d_{t_q}}}$ are left out. $\lhd$

Apart from the x_is and y_is visible in the formula, the x_i variables also occur in the αs, the y_is occur in the βs, and the z_is in the γs.

Remark. Here, 'recursive' refers to the fact that the definition of

$$\phi_{S_T(X_3^{q,t_q})}\left(z_1, \ldots, z_{d_{t_q}}\right)$$

is necessary to define the formula $\phi_T(x_1, \ldots, x_d)$. The word 'inductive' may also be used. $\lhd$

$$
\begin{tikzcd}
 & & A & \\
T(\bot = X_1) \arrow[rru, bend left, "f_1"] \arrow[r, "T(m_2^q)"] & T(X_2^q) \arrow[ru, bend left, "e_2"] \arrow[r, "T(m_3^{q,t_q})"] & T(X_3^{q,t_q}) \arrow[u, "f_3"'] \arrow[r, "\dots"] & \cdots \arrow[lu, bend right, "\dots"']
\end{tikzcd}
$$

The interpretation of each fragment of the formula in terms of the tree is as follows. (The formula piece is given first, then its meaning).

$\forall_{x_1,\dots,x_d \in A}$

For every **map** $f : T(\bot) \to A$...

$$\bigwedge_{\rho,\bar{x}} \alpha_{\rho,\bar{x}}$$

The map f_1 is a morphism.

$$\bigvee_{q} \left[\exists_{y_1,\dots,y_{n_q}} (\cdots)\right]$$

There is some X_2^q and map $e_2 : T(X_2^q) \to A$.

$$\bigwedge_{\rho,\bar{y}} \beta^q_{\rho,\bar{y}}$$

The map e_2 is a morphism.

$$\bigwedge_{r_q} \left(x_{i_{r_q}} = y_{j_{r_q}}\right)$$

The commutativity property $e_2 \circ T(m_2^q) = f_1$ holds.

$$\bigwedge_{t_q} \forall_{z_1,\dots,z_{d_{t_q}}} [\cdots]$$

For each X_3^{q,t_q} (with morphism from X_2^q) and for every map $f : T(X_3^{q,t_q}) \to A$...

$$\bigwedge_{s_{t_q}} \left(y_{i_{s_{t_q}}} = z_{j_{s_{t_q}}} \right)$$

The commutativity property $f_3 \circ T(m_3^{q,t_q}) = e_2$ holds.

$$\bigwedge_{t_q,\rho,\bar{z}} \gamma_{\rho,\bar{z}}^{q,t_q}$$

The map f_3 is a morphism.

$$\phi_{S_T(X_3^{q,t_q})} \left(z_1, \ldots, z_{d_{t_q}} \right)$$

The subtree $S_T(X_3^{q,t_q})$ is injective under the assignment f_3.

6.4.3 Formal Definition

Definition 6.4.3.1. Fix a finite signature $\mathbf{\Sigma}$ and let **RStr** be the category of relational structures in this signature.

Let $T : \mathbb{T} \to \mathbf{RStr}$, whose image consists of **finite** relational structures.

Let $x_1, \ldots, x_d \in T(\perp)$ be exactly the elements of $T(\perp = X_1)$.

Given the definition of the formula $\phi_{S_T(X_3^{q,t_q})} \left(z_1, \ldots, z_{d_{t_q}} \right)$ for each subtree $S_T(X_3^{q,t_q})$, the formula $\phi_T(x_1, \ldots, x_d)$ is defined recursively by:

$$\bigwedge_{\rho,\bar{x}} \alpha_{\rho,\bar{x}} \Rightarrow \bigvee_{q} \left[\exists_{y_1,\ldots,y_{n_q}} \left(\bigwedge_{\rho,\bar{y}} \beta_{\rho,\bar{y}}^{q} \wedge \bigwedge_{i} (x_i = y_{j_i}) \wedge \bigwedge_{t_q} \delta_{t_q} \right) \right]$$

where $\delta_{t_q} \equiv$

$$\forall_{z_1,\ldots,z_{d_{t_q}}} \left[\left(\bigwedge_{i} (y_i = z_{j_i}) \wedge \bigwedge_{\rho,\bar{z}} \gamma_{\rho,\bar{z}}^{q,t_q} \right) \quad \Rightarrow \quad \phi_{S_T(X_3^{q,t_q})} \left(z_1, \ldots, z_{d_{t_q}} \right) \right].$$

For each relation symbol ρ, number n_ρ such that $\rho \in \mathbf{\Sigma}_{n_\rho}$, and tuple $\bar{x} \in T(\perp)^{n_\rho}$, if

$$\bar{x} = \langle x_{k_1}, \ldots, x_{k_{n_\rho}} \rangle \in R_{n_\rho}^{T(\perp)}(\rho)$$

then define

$$\alpha_{\rho,\bar{x}} \quad \equiv \quad \rho(x_{k_1}, \ldots, x_{k_{n_\rho}}),$$

otherwise define

$$\alpha_{\rho,\bar{x}} \quad \equiv \quad \mathbf{True}.$$

Remark. The conjunction could be viewed as

$$\bigwedge_{\bar{x}\in R_{n_\rho}^{T(\perp)}(\rho)} \alpha_{\rho,\bar{x}}.$$

(in which case it would not matter what $\alpha_{\rho,\bar{x}}$ is equal to outside of the given cases.)

Instead, for all the $\rho, \bar{x}$ that do not satisfy the given condition, $\alpha_{\rho,\bar{x}}$ is declared to be tautologically true so that it does not have an effect on the conjunction.

The subformulae β and γ are treated similarly. $\lhd$

Remark. The number n_ρ may be zero, in which case $\bar{x}$ is an empty tuple and the αs are all sentences. $\lhd$

For any relation symbol ρ, number q representing the X in this layer of the tree, number n_ρ representing the free variables in the relational symbol ρ, and tuple of variables $\bar{y} \in T(X_1^q)^{n_\rho}$, if

$$\bar{y} = \langle y_{k_1}, \ldots, y_{k_{n_\rho}} \rangle \in R_{n_\rho}^{T(X_1^q)}(\rho)$$

then define

$$\beta^q_{\rho,\bar{y}} \quad \equiv \quad \rho(y_{k_1}, \ldots, y_{k_{n_\rho}}),$$

otherwise

$$\beta^q_{\rho,\bar{y}} \quad \equiv \quad \textbf{True}.$$

For each layer two object X_2^q with corresponding morphism $m_2^q : \perp \to X_2^q$, if $T(X_2^q) = \{y_1, y_2, \ldots\}$, then for each pair i, j such that

$$T(m_2^q)(x_i) = y_j$$

define $j_i = j$ and declare

$$x_i = y_{j_i}.$$

Similarly, for each level three object X_3^{q,t_q} with corresponding morphism m_3^{q,t_q} : $X_3^q \to X_3^{q,t_q}$, and for each pair i, j such that

$$z_j = T(m_3^{q,t_q})(y_i)$$

define $j_i = j$ and declare

$$y_i = z_{j_i}.$$

Remark. If there are no xs or ys (because some structures in the tree are empty), then the corresponding conjunctions are left out. $\lhd$

For each ρ, with $\rho \in \Sigma_{n_\rho}$, and $\bar{z} \in T(X_3^{q,t_q})^{n_\rho}$, if

$$\bar{z} = \langle z_{k_1}, \dots, z_{k_{n_\rho}} \rangle \in R_{n_\rho}^{T(X_3^{q,t_q})}(\rho)$$

then define

$$\gamma_{\rho,\bar{z}}^{q,t_q} \quad \equiv \quad \rho(z_{k_1}, \dots, z_{k_{n_\rho}}).$$

else

$$\gamma_{\rho,\bar{z}}^{q,t_q} \quad \equiv \quad \textbf{True}.$$

Altogether, this gives:

$$\begin{aligned}
&\bigwedge_{\bar{x} \in R_{n_\rho}^{T(\perp)}(\rho)} \rho(x_1, \dots, x_{n_\rho}) \quad \Rightarrow \\
&\bigvee_q [\exists_{y_1,\dots,y_{n_q}} (\bigwedge_{\bar{y} \in R_{n_\rho}^{T(X_1^q)}(\rho)} \rho(y_1, \dots, y_{n_\rho}) \wedge \bigwedge_i (x_i = y_{j_i}) \wedge \\
&\quad \bigwedge_{t_q} [\forall_{z_1,\dots,z_{t_q}} (\bigwedge_i (y_i = z_{j_i}) \wedge \bigwedge_{\bar{z} \in R_{n_\rho}^{T(X_3^{q,t_q})}(\rho)} \rho(z_{k_1}, \dots, z_{k_{n_\rho}}) \\
&\qquad \Rightarrow \quad \phi_{S_T(X_3^{q,t_q})}(z_1, \dots, z_{t_q}))])].
\end{aligned}$$

$\triangleleft$

6.4.4 Theorem: Tree Injective iff Sentence True

Theorem 6.4.4.1. Given a tree T, and an object $A \in \mathbf{RStr}$, then

$$\forall_{x_1,\dots,x_d \in A} \phi_T(x_1, \dots, x_d) \qquad \text{iff} \qquad A \models T.$$

Proof. A map $f : T(\perp) \to A$ is a morphism if the following property is satisfied:

For every relational symbol ρ, if the interpretation in $T(\perp)$ holds then the interpretation holds for the images under f in A as well, i.e.,

$$\langle x_{k_1}, \dots, x_{k_{n_\rho}} \rangle \in R_{n_\rho}^{T(\perp)}(\rho) \quad \Rightarrow \quad \langle f(x_{k_1}), \dots, f(x_{k_{n_\rho}}) \rangle \in R_{n_\rho}^{\mathcal{A}}(\rho).$$

Now, if every f fails to be a morphism, then for each f there is some ρ_f such that

$$\langle x_{k_1}, \dots, x_{k_{n_{\rho_f}}} \rangle \in R_{n_{\rho_f}}^{T(\perp)}(\rho_f)$$

holds, but not

$$\langle f(x_{k_1}), \dots, f(x_{k_{n_{\rho_f}}}) \rangle \in R_{n_{\rho_f}}^{\mathcal{A}}(\rho_f).$$

Then, by definition of the formula ϕ, for each assignment $f(x_1), \ldots, f(x_d)$ of variables, there is some ρ such that, for some tuple $\langle x_{k_1}, \ldots, x_{k_{n_\rho}} \rangle \in R^{\mathcal{A}}_{n_\rho}(\rho_f)$, the statement

$$\rho(x_{k_1}, \ldots, x_{k_{n_\rho}})$$

is false for A under the assignment f, and so

$$\bigwedge_{\rho,\bar{x}} \alpha_{\rho,\bar{x}} \qquad \text{i.e.,} \qquad \bigwedge_{\rho,\bar{x} \in R^{T(\perp)}_{n_\rho}(\rho)} \rho(f(x_{k_1}), \ldots, f(x_{k_{n_\rho}}))$$

is false under the interpretation f, so the implication $\wedge_{\rho,\bar{x}} \alpha_\rho \Rightarrow \cdots$ is trivially true and so $\forall_{x_1,\ldots,x_d}[\phi(x_1, \ldots, x_d)]$ is true.

Remark. It may be a bit confusing why here $\alpha_{\rho,\bar{x}}$ is $\rho(f(x_{k_1}), \ldots, f(x_{k_{n_\rho}}))$ instead of $\rho(x_{k_1}, \ldots, x_{k_{n_\rho}})$. The elements x fulfil a dual role as elements of $T(\perp)$ and as variables speaking about elements of A. However, for variables to be interpreted as specific elements of A is to give an assignment f.

To check the truth of $\forall_{x_1,\ldots,x_d}\phi(x_1, \ldots, x_d)$ is to check its truth for each assignment $f(x_1), \ldots, f(x_d)$. The formulae are equivalent

$$\forall_{x_1,\ldots,x_d}\phi(x_1, \ldots, x_d) \quad \equiv \quad \forall_f \phi(f(x_1), \ldots, f(x_d))$$

where in the first case the $x_1, \ldots, x_d$ are variables naming elements of A, whereas in the second case they are the elements of $T(\perp)$ and each $f : T(\perp) \to A$ is an arbitrary map, not necessarily a morphism.

Hence the statement $\rho(x_{k_1}, \ldots, x_{n_\rho})$ interpreted in A is $\rho(f(x_{k_1}), \ldots, f(x_{n_\rho}))$ for some assignment f. $\triangleleft$

($\Rightarrow$) The proof is inductive on the size of the tree. Assume it holds for all subtrees of T of the form $S_T(X_3^{q,t_q})$.

Assume that there is some **morphism** $f : T(\perp) \to A$. Let $f_1 : T(\perp) \to A$ be an arbitrary such morphism. It is shown that $A \models T[f_1]$.

Since f_1 is a morphism, for every symbol ρ, the tuple

$$\langle f_1(x_{k_1}), \ldots, f_1(x_{k_{n_\rho}}) \rangle \in R^{\mathcal{A}}_{n_\rho}(\rho)$$

(is related by $R^{\mathcal{A}}_{n_\rho}(\rho)$ in A) if $\langle x_{k_1}, x_{k_{n_\rho}} \rangle \in R^{T(\perp)}_{n_\rho}(\rho)$ and so by definition of ϕ, the subformula

$$\bigwedge_{\rho,\bar{x}} \alpha_{\rho,\bar{x}} \qquad \text{i.e.,} \qquad \bigwedge_{\rho,\bar{x} \in R^{T(\perp)}_{n_\rho}(\rho)} \rho(f_1(x_{k_1}), \ldots, f_1(x_{k_{n_\rho}}))$$

is true in A.

Hence, since ϕ is true for every assignment by a morphism f_1, and since it asserts

$$\bigwedge_{\rho,\bar{x}} \alpha_{\rho,\bar{x}} \Rightarrow \bigvee_q \left[\exists_{y_1,\ldots,y_{n_q}} \left(\bigwedge_{\rho,\bar{y}} \beta^q_{\rho,\bar{y}} \wedge \bigwedge_{r_q} \left(x_{i_{r_q}} = y_{j_{r_q}}\right) \wedge \bigwedge_{t_q} \delta_{t_q}\right)\right]$$

then by modus ponens the following disjunction holds for the assignment f_1

$$\bigvee_q \left[\exists_{y_1,\ldots,y_{n_q}\in A} \left(\bigwedge_{\rho,\bar{y}} \beta^q_{\rho,\bar{y}} \wedge \bigwedge_{r_q} \left(f_1(x_{i_{r_q}}) = y_{j_{r_q}}\right) \wedge \bigwedge_{t_q} \delta_{t_q}\right)\right].$$

So for some q the inside holds. Then consider the layer two object X^q_2 with corresponding morphism $m^q_2 : \perp \to X^q_2$.

Define a **map** $e_2 : T(X^q_2) \to A$ by mapping $y_1, \ldots, y_{n_q} \in T(X^q_2)$ to the elements $y_1, \ldots, y_{n_q} \in A$ asserted by the existential quantifier in the above statement.

The subformula

$$\bigwedge_{\rho,\bar{y}} \beta^q_{\rho,\bar{y}} \qquad \text{i.e.,} \qquad \bigwedge_{\rho,\bar{y}\in R^{T(X^q_2)}_{n_\rho}(\rho)} \rho(e_2(y_{k_1}), \ldots, e_2(y_{k_{n_\rho}}))$$

ensures that any relation holding in $T(X^q_2)$ also holds in A and so the map e_2 is actually a morphism.

The subformula

$$\bigwedge_{r_q} (f_1(x_{i_{r_q}}) = e_2(y_{j_{r_q}}))$$

means that

$$f_1(x_{i_{r_q}}) = e_2(y_{j_{r_q}}) = e_2(T(m^q_2)(x_{i_{r_q}}))$$

and so $e_2 \circ T(m^q_2) = f_1$.

Finally, given any object X^{q,t_q}_3, morphism $m^{q,t_q}_3 : X^q_2 \to X^{q,t_q}_3$, and morphism $f_3 : T(X^{q,t_q}_3) \to A$ which maps each $z_1, \ldots, z_{d_{t_q}} \in T(X^{q,t_q}_3)$ somewhere in A, then, as above, the formula

$$\bigwedge_{t_q} \delta_{t_q} \quad \text{i.e.,}$$

$$\bigwedge_{t_q} \forall_{z_1,\ldots,z_{d_{t_q}}} \left[\left(\bigwedge_{s_{t_q}} \left(y_{i_{s_{t_q}}} = z_{j_{s_{t_q}}}\right) \wedge \bigwedge_{\rho,\bar{z}\in R^{T(X^{q,t_q}_3)}_{n_\rho}(\rho)} \rho(z_{k_1}, \ldots, z_{k_{n_\rho}})\right) \Rightarrow \phi_{S_T(X^{q,t_q}_3)}(z_1, \ldots, z_{d_{t_q}}))\right]$$

ensures that, provided $f_3 \circ T(m_3^{q,t_q}) = e_2$, then $\phi_{S_T(X_3^{q,t_q})}(z_1, \ldots, z_{t_q})$ holds and so by the inductive hypothesis, $A \models S_T(X_3^{q,t_q})$.

($\Leftarrow$) Assume $A \models T$.

For any $a_1, \ldots, a_d \in A$ there is corresponding map $f_1 : T(\perp) \to A$ which takes the elements $x_1, \ldots, x_d \in T(\perp)$ to the elements $a_1, \ldots, a_d \in A$. Or else, if $T(\perp)$ is empty, then there is a unique $f_1 : T(\perp) \to A$. If A is empty, there might not be a map from $T(\perp)$.

If f_1 is not a homomorphism, then there is some relation symbol ρ such that

$$\langle x_{k_1}, \ldots, x_{k_{n_\rho}} \rangle \in R_{n_\rho}^{T(\perp)}(\rho),$$

but not

$$\langle a_{k_1}, \ldots, a_{k_{n_\rho}} \rangle \in R_{n_\rho}^{\mathcal{A}}(\rho).$$

Then, by definition of the formula ϕ,

$$\rho(f_1(x_{k_1}), \ldots, f_1(x_{k_{n_\rho}}))$$

is false, and so

$$\bigwedge_{\rho,\bar{x}} \alpha_\rho \qquad \text{i.e.,} \qquad \bigwedge_{\rho,\bar{x} \in R_{n_\rho}^{T(\perp)}(\rho)} \rho(f(x_{k_1}), \ldots, f(x_{k_{n_\rho}}))$$

is false so the implication $\wedge_{\rho,\bar{x}}(\alpha_{\rho,\bar{x}}) \Rightarrow \cdots$ is trivially true and so the formula $\phi(f_1(x_1), \ldots, f_1(a_d))$ is true.

Remark. In particular, if ρ is a nullary relational symbol then the empty tuple $\langle\rangle$ may be in $R_{n_\rho}^{T(\perp)}(\rho)$ but not in $R_{n_\rho}^{\mathcal{A}}(\rho)$. This corresponds to some proposition being true of $T(\perp)$ but not $\mathcal{A}$. $\lhd$

Assume now that f_1 is a morphism of relational structures.

Since A is injective with respect to the tree, there is X_2^q, morphism $m_2^q : \perp \to X_2^q$ and morphism $e_2 : T(X_2^q) \to A$ with $e_2 \circ T(m_2^q) = f_1$.

Then there are elements in A given by the images under e_2 of $y_1, \ldots, y_{n_q} \in T(X_2^q)$. Since e_2 is a morphism of relational structures, then

$$\langle e_2(y_{k_1}), \ldots, e_2(y_{k_{n_\rho}}) \rangle \in R_{n_\rho}^{\mathcal{A}}(\rho)$$

are related in A whenever $\langle y_{k_1}, \ldots, y_{k_{n_\rho}} \rangle \in R_{n_\rho}^{T(X_2^q)}(\rho)$ and so

$$\bigwedge_{\rho,\bar{y}} \beta_{\rho,\bar{y}}^q \qquad \text{i.e.,} \qquad \bigwedge_{\rho,\bar{y} \in R_{n_\rho}^{T(X_2^q)}(\rho)} \rho(e_2(y_{k_1}), \ldots, e_2(y_{k_{n_\rho}}))$$

is true in A.

Remark. If A is empty, then there are no elements in A given by the images under e, but then also $T(X_2^q)$ must itself be empty by the fact that A is injective with respect to T. $\triangleleft$

Since $e_2 \circ T(m_2^q) = f_1$, then

$$\bigwedge_{r_q} (f_1(x_{i_{r_q}}) = e_2(y_{j_{r_q}})).$$

Finally, given t_q and any $z_1, \ldots, z_{d_{t_q}}$ there is object X_3^{q,t_q}, morphism $m_3^{q,t_q} : X_2^q \to X_3^{q,t_q}$, and map $f_3 : T(X_3^{q,t_q}) \to A$ which takes $z_1, \ldots, z_{d_{t_q}} \in T(X_3^{q,t_q})$ to $f_3(z_1), \ldots, f_3(z_{d_{t_q}}) \in A$.

Assume the hypothesis of the implication of δ_{t_q}, i.e., assume

$$\left(\bigwedge_{s_{t_q}} \left(e_2(y_{i_{s_{t_q}}}) = f_3(z_{j_{s_{t_q}}}) \right) \wedge \bigwedge_{\rho,\, \bar{z} \in R_{n_\rho}^{T(X_3^{q,t_q})}(\rho)} \rho(f_3(z_{k_1}), \ldots, f_3(z_{k_{n_\rho}})) \right).$$

If

$$\langle z_{k_1}, \ldots, z_{k_{n_\rho}} \rangle \in R_{n_\rho}^{T(X_3^{q,t_q})}(\rho)$$

then

$$\langle f_3(z_{k_1}), \ldots, f_3(z_{k_{n_\rho}}) \rangle \in R_{n_\rho}^{\mathcal{A}}(\rho)$$

and so f_3 is a relational structure morphism.

Since

$$\bigwedge_{s_{t_q}} (e_2(y_{i_{s_{t_q}}}) = f_3(z_{j_{s_{t_q}}}))$$

holds, then $f_3 \circ T(m_3^{q,t_q}) = e_2$.

By the definition of injectivity of trees, $A \models S_T(X_3^{q,t_q})$.

Hence, by the inductive hypothesis $\phi_{S_T(X_3^{q,t_q})}(f_3(z_1), \ldots, f_3(z_{t_q}))$ holds.

Then the conclusion of the implication holds.

So

$$\bigwedge_{t_q} \delta_{t_q} \quad \text{i.e.,}$$

$$\bigwedge_{t_q} \forall_{z_1,\ldots,z_{d_{t_q}}} \left[\left(\bigwedge_{s_{t_q}} \left(e_2(y_{i_{s_{t_q}}}) = z_{j_{s_{t_q}}} \right) \wedge \bigwedge_{\rho, \bar{z} \in R_{n_\rho}^{T(X_3^{q,t_q})}(\rho)} \rho(z_{k_1}, \ldots, z_{k_{n_\rho}}) \right) \Rightarrow \phi_{S_T(X_3^{q,t_q})}(z_1, \ldots, z_{d_{t_q}}) \right]$$

holds. □

6.5 Trees of Relational Structures from First-Order Formulae

6.5.1 Preliminaries

The results of the following section still take place within the category of relational structures.

Given a specific quantifier-free formula $\phi(x_1, \ldots, x_d)$, it is possible to construct a tree T_ϕ (with even $k > d$) such that, for any object $A \in \mathbb{C} = \mathbf{RStr}$,

$$\forall_{x_1}\exists_{x_2}\cdots,\forall_{x_{k-1}}\exists_{x_k}[\phi(x_1,\ldots,x_d)] \qquad \text{iff} \qquad A \models T_\phi.$$

This is the main result of the paper [AN79].

A summary of the method is provided below.

It is shown that it is possible to form a tree given that ϕ is of the form

$$\bigwedge_{1\leq i\leq z}\left(\alpha_i \to \bigvee_{1\leq j\leq p_i} \beta^i_j\right)$$

where the α_i are conjunctions of atomic formulae and β^i_j are atomic formulae. By [HMT71], this is equivalent to any propositional formula, and so any formula can be represented.

Definition 6.5.1.1. Given a conjunction of atomic formulae α, define $\equiv_\alpha$ to be the equivalence relation generated by

$$\{\langle x_i, x_j\rangle \mid \text{“}x_i = x_j\text{” occurs in } \alpha\}.$$

(In other words, $\equiv_\alpha$ is the smallest equivalence relation which is a superset of the above set). ◁

Definition 6.5.1.2. Consider a conjunction of atomic formulae

$$\alpha = \bigwedge_{1\leq i\leq n} \gamma_i$$

Each γ_i is either a relational symbol $\rho(x_{j_1}, \ldots, x_{j_m})$ or an equality statement $x_i = x_j$ for some x_i and x_j.

Let m be a number at least as large as the largest index of a free variable occurring in the conjunction. In other words,

$$m \geq \max\{j \mid x_j \text{ occurs in } \gamma_i \text{ for some } i \leq n\}.$$

Define $B'_m = \{x_1, \ldots, x_m\}$.

For each $n \in \mathbb{N}_0$ define $(R^\alpha_n)' : \Sigma_n \to \mathcal{P}\left((B'_m)^n\right)$ on each $\rho \in \Sigma_n$ by

$$(R^\alpha_n)'(\rho) = \left\{\langle x_{j_1}, \ldots, x_{j_n}\rangle \in \left(B'_m\right)^n \mid \rho(x_{j_1}, \ldots, x_{j_n}) \text{ occurs in } \alpha\right\}.$$

(More precisely, 'occurs in' means 'is a subformula of'.)

Define $b'_m(\alpha) = \langle B'_m, \langle (R^\alpha_n)'\rangle_{n\in\mathbb{N}}\rangle$, where $\equiv_\alpha$ is now interpreted as a relation on B'_m.

Now define $B_m = B'_m / \equiv_\alpha$.

Denote the equivalence class of $x_i \in B'_m$ in B_m by $[x_i]_\alpha$.

For each $n \in \mathbb{N}_0$ define $R^\alpha_n : \Sigma_n \to \mathcal{P}\left((B_m)^n\right)$ on each $\rho \in \Sigma_n$ by

$$\begin{aligned} R^\alpha_n(\rho) = \{ & \langle [x_{j_1}]_\alpha, \ldots, [x_{j_n}]_\alpha\rangle \in (B_m)^n \mid \\ & \langle x_{j_1}, \ldots, x_{j_n}\rangle \in (R^\alpha_n)'(\rho), \\ & \text{for some tuple of representatives } \langle x_{j_1}, \ldots, x_{j_n}\rangle\}. \end{aligned}$$

Remark. At first glance, it may seem that it must still be shown that the definition is well-defined, regardless of the choice of representatives of the equivalence classes. In fact, it is automatically well-defined because there are no choices of representatives made.

Note that this definition states that if there is **any** tuple of choices of representatives which are related, then the equivalence classes are related (even if, say, all other choices of representatives are not related). ◁

Then the m**-element model of** α, denoted $b_m(\alpha)$, is defined by

$$b_m(\alpha) = \langle B_m, \langle R^\alpha_n\rangle_{n\in\mathbb{N}_0}\rangle .$$

◁

Remark. Even though it is called the m-element model, b_m won't in general have m elements (although b'_m will). ◁

The motivation for this is that it is essentially the most efficient way of forming a relational structure in which α is true. Most efficient in the sense that there are no more true relations than those implied by α.

6.5.2 Formula

Assume ϕ is of the form

$$\bigwedge_{i\leq z}\left(\alpha_i \to \bigvee_{j\leq p_i} \beta^i_j\right)$$

where $z \in \mathbb{N}_0$ and each $p_i \in \mathbb{N}_0$, such that each α_i is a conjunction of atomic formulae, and each β^i_j is an atomic formula.

Remark. If $z = 0$, the formula is interpreted as **True**, i.e., tautologically true. If some $p_i = 0$, then that disjunction is interpreted as **False**, i.e., tautologically false. This allows, e.g., $\neg\alpha_i$ by $\alpha_i \rightarrow$ **False**. $\triangleleft$

Fix $n \in \mathbb{N}^+$ and assume also that the largest index of a variable occurring in ϕ is at most $2n$. So

$$\max\{i \mid x_i \text{ occurs in } \phi\} \leq 2n.$$

Define $k = 2n$.

6.5.3 Construction of Tree

We are now ready to construct the tree.

Definition 6.5.3.1. The **tree category for** ϕ, denoted $\mathbb{T}_\phi$, is formed below:

The objects are strings of numbers of one of the following forms:

- One to k zeroes. $0^g = 00\ldots0$ (g zeroes with $l \leq g \leq k$). Let $\perp = 0$.
- A string of k zeroes followed by a number $1 \leq i \leq z$, denoted $0^k i$.
- A string of k zeroes followed by a number $1 \leq i \leq z$ followed by a number $1 \leq j \leq p_i$, denoted $0^k ij$.

The morphisms are simply arrows. There are morphisms precisely in the following cases:

- $\bar{m}^{g+1} : 0^g \rightarrow 0^{g+1}$ for each $1 \leq g < k$.
- $m^i : 0^k \rightarrow 0^k i$ for each $1 \leq i \leq z$.
- $m^i_j : 0^k i \rightarrow 0^k ij$ for each $1 \leq i \leq z$ and $1 \leq j \leq p_i$.
- Identity morphisms.
- All morphisms formed from compositions of such morphisms.

$\triangleleft$

$$\begin{tikzcd}
 & & & & 0^k 11 & \\
 & & & & \vdots & \\
 & & & 0^k 1 \arrow[uur, "m^1_1"] \arrow[ur] \arrow[r, "m^1_{p_1}"'] & 0^k 1 p_1 & 0^k i 1 \\
 & & & \vdots & & \vdots \\
\perp = 0 \arrow[r, "\bar{m}^1"] & \cdots \arrow[r, "\bar{m}^k"] & 0^k \arrow[uur, "m^1"] \arrow[ur] \arrow[r, "m^i"] \arrow[dr] \arrow[ddr, "m^z"'] & 0^k i \arrow[uurr, "m^i_1"] \arrow[urr] \arrow[rr, "m^i_j"] \arrow[drr] \arrow[ddrr, "m^i_{p_i}"'] & & 0^k ij \\
 & & & \vdots & & \vdots \\
 & & & 0^k z \arrow[r, "m^z_1"] \arrow[dr] \arrow[ddr, "m^z_{p_z}"'] & 0^k z 1 & 0^k i p_i \\
 & & & & \vdots & \\
 & & & & 0^k z p_z &
\end{tikzcd}$$

Definition 6.5.3.2. $R^{\emptyset}_l$ is defined (for any set A) for each $l \in \mathbb{N}$ as follows:

$$R^{\emptyset}_l : \Sigma_l \to \mathcal{P}\left(A^l\right) : \rho \mapsto \emptyset.$$

Define the **trivial relational family** as

$$R^{\emptyset} = \left\langle R^{\emptyset}_l \right\rangle_{l \in \mathbb{N}}.$$

This corresponds to interpreting every relational symbol as an empty relation.

$\triangleleft$

Definition 6.5.3.3. The **tree for** ϕ is a functor $T : \mathbb{T}_\phi \to \mathbf{RStr}$ determined by the following:

1. $T(0^g) = \left\langle \{x_i \mid 1 \leq i \leq g\}, (R^{\emptyset})^{\{x_i \mid 1 \leq i \leq g\}} \right\rangle$, for every $1 \leq g \leq k$.
2. $T(0^k i) = b_k(\alpha_i)$ for every $1 \leq i \leq z$.

3. $T(0^k ij) = b_k\left(\alpha_i \wedge \beta^i_j\right)$ for every $1 \leq i \leq z$ and $1 \leq j \leq p_i$.
4. $T(\bar{m}^{g+1} : 0^g \to 0^{g+1})(x_l) = x_l$ for every $1 \leq g \leq k$ and $1 \leq l \leq g$.
5. $T(m^i : 0^k \to 0^k i)(x_l) = [x_l]_{\alpha_i}$ for every $1 \leq i \leq z$ and $1 \leq l \leq k$.
6. $T(m^i_j : 0^k i \to 0^k ij)([x_l]_{\alpha_i}) = ([x_l]_{\alpha_i \wedge \beta^i_j})$ for every $1 \leq i \leq z$, $1 \leq j \leq p_i$ and $1 \leq l \leq k$.

$\triangleleft$

In what follows, the elements of $T(0^g)$ will be labelled $x^g_1, \ldots, x^g_g$. However, when there is no ambiguity they may simply be labelled $x_1, \ldots, x_g$.

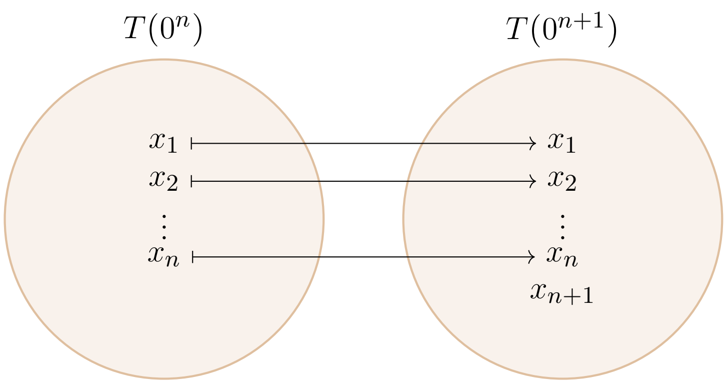
6.5.4 Intuition for Definition

The intuition for the above definition is described first, and the formal proof follows.

The first part of the tree (the non-branching part), corresponds to the quantifiers.

$$
\begin{tikzcd}
 & & A & & \\
T(0) \arrow[urr, bend left, "\forall x_1"] \arrow[r, hook, "x_l \mapsto x_l"] & T(00) \arrow[ur, "\exists x_2"] \arrow[r, hook, "x_l \mapsto x_l"] & T(000) \arrow[u, "\forall x_3"'] \arrow[r, hook, "x_l \mapsto x_l"] & T(0000) \arrow[ul, "\exists x_4"'] \arrow[r] & \cdots \arrow[ull, bend right]
\end{tikzcd}
$$

Each of the relational structures $(T(0), T(00), \ldots)$ in the image of this part of the tree has no relations. Hence the only sentence-fragments for A represented by this part of the tree do not refer to relations. All of the morphisms here are injective, and hence no different elements are collapsed.



The object $T(0)$ has a single element, $T(00)$ has two elements and so forth. Each successive object has an extra element, with the previous object as a subset.

Consider the definition of injectivity of a tree (Definition 6.3.3.1). The properties of injectivity need to be true for every morphism from $T(0)$ to A and this corresponds to the subformula being true for every $x_1 \in A$.

Similarly, the properties for injectivity need to be true for some morphism from $T(00)$ to A. Since the diagram needs to commute, x_1 is already determined, and so the morphism only declares that some x_2 needs to exist.

This process continues and so one obtains that the sentence begins with

$$\forall_{x_1}\exists_{x_2}\cdots\forall x_{k-1}\exists_{x_k}.$$

Remark. Technically, according to Definition 6.4.3.1, there are quantifiers named for each object and they are collapsed by equality to match the above statement.

More precisely, if the image of the tree has two objects $T(0)$ and $T(00)$, the corresponding sentence fragment is:

$$\forall_{x_1^1}\exists_{x_1^2,x_2^2}(x_1^1 = x_1^2)$$

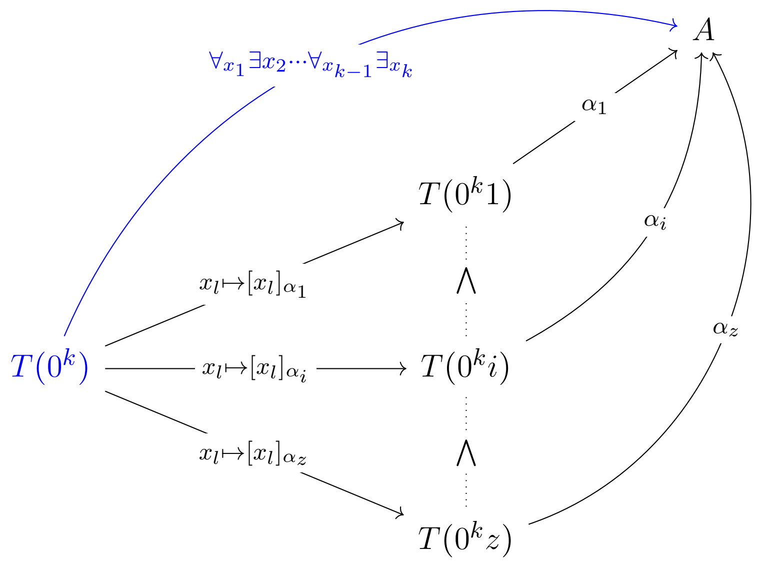
Similarly, if the tree has four objects $T(0), \ldots, T(0000)$, the corresponding sentence fragment is:

$$\begin{aligned}
\forall_{x_1^1}\exists_{x_1^2,x_2^2}\forall_{x_1^3,x_2^3,x_3^3}\exists_{x_1^4,x_2^4,x_3^4,x_4^4}\\
& (\quad x_1^1 = x_1^2 \\
& \wedge \quad x_1^2 = x_1^3 \quad \wedge \quad x_2^2 = x_2^3 \\
& \wedge \quad x_1^3 = x_1^4 \quad \wedge \quad x_2^3 = x_2^4 \quad \wedge \quad x_3^3 = x_3^4).
\end{aligned}$$

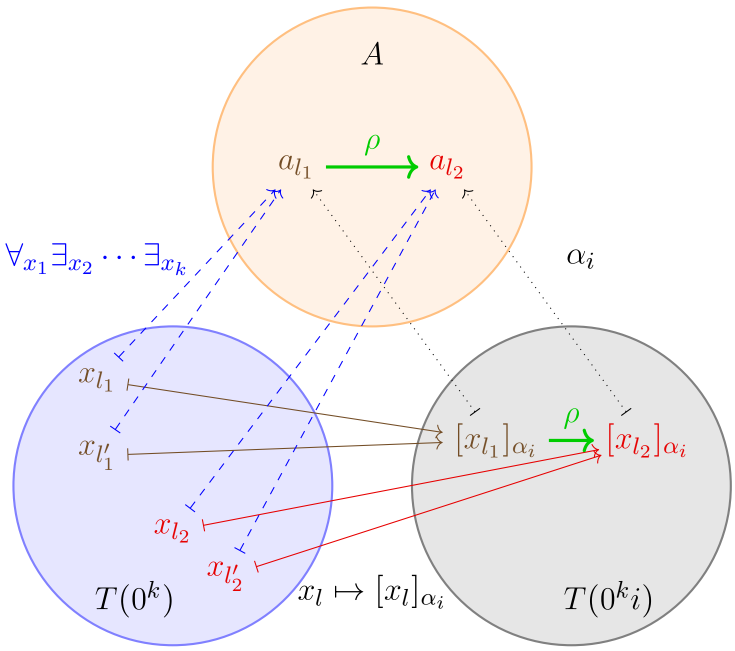
However, this will be equivalent to the corresponding sentence starting with the quantifiers over only single variables. ◁

The subtree starting at $T(0^k)$ corresponds to the formula ϕ.

In the diagram below, each morphism to A has been labelled with the part of the formula for which it is most important. The $\wedge$ symbols on the vertical dotted lines represent conjunction.



Note that the morphisms going from $T(0^k)$ into each of the $T(0^ki)$ correspond to the equality symbols of α_i, whereas the morphisms going from each $T(0^ki)$ into A correspond to the relational symbols of α_i.



In the diagram above, note how the quantifier map corresponds to arbitrary selection of elements of A, since there are no relation symbols. Note that the $x_l \mapsto [x_l]_{\alpha_i}$ forces the images of x_l in A under the quantifier map to be equal since the quantifier map has to commute with the map labelled α_i.

Note also that since the map labelled α_i is a morphism, then also every relation ρ holding in $T(0^k i)$ must hold for the images under the morphism labelled α_i. Hence any ρ which is a subformula of α_i is forced to hold in this way in A.

Each of the morphisms $T(m^i)$ along with the subtree at $T(0^k i)$ corresponds to the fragment

$$\alpha_i \to \bigvee_{j \leq p_i} \beta^i_j$$

of the conjunction

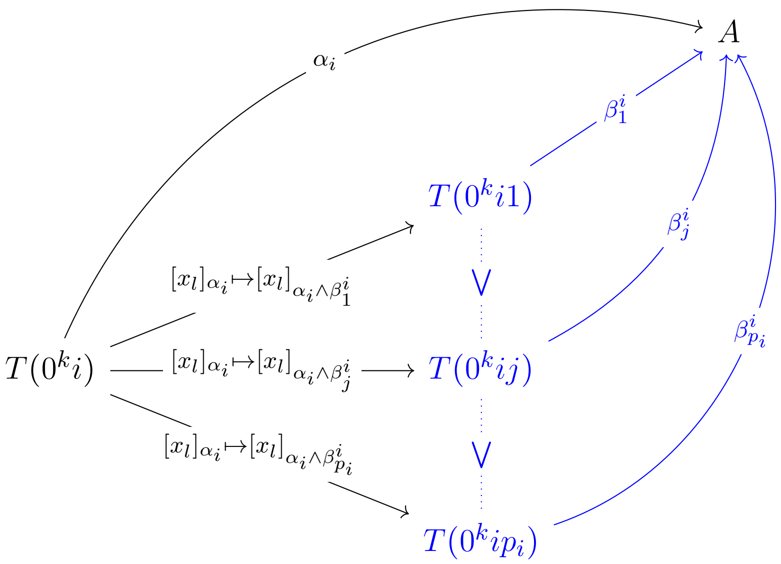
$$\bigwedge_{i \leq z} \left(\alpha_i \to \bigvee_{j \leq p_i} \beta^i_j \right).$$

Now consider a specific i, and the subtree starting at $T(0^k i)$.

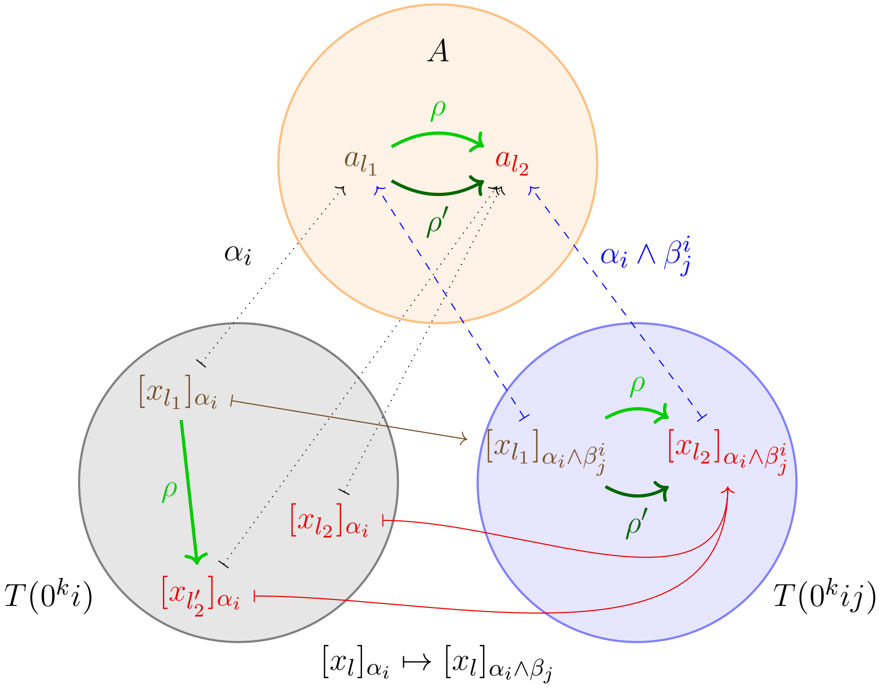
Each of the morphisms $T(m^i_j)$ along with the object $T(0^k ij)$ corresponds to the atomic fragment β^i_j of the disjunction

$$\bigvee_{j \leq p_i} \beta^i_j.$$

In the following diagram, the $\vee$ symbols on the vertical dotted lines represent conjunction.



Again the morphisms going from $T(0^k i)$ into each of the $T(0^k ij)$ correspond to the equality symbols of β^i_j, whereas the morphisms going out of each $T(0^k ij)$ into A correspond to the relational symbols of β^i_j.



In the diagram above, note again how the morphisms $[x_l]_{\alpha_i} \mapsto [x_l]_{\alpha_i \wedge \beta^i_{p_i}}$ force elements in A (already forced to be equal by equations in α_i) to be equal by equations in β^i_j.

Further, note that the morphism labelled $\alpha_i \wedge \beta^i_j$ forces relations ρ' which are a subformula of β^i_j to hold in A.

Hence, if the relations given by α_i hold, then the morphism labelled α_i exists (with appropriate commutativity properties), and this forces the morphism labelled by $\alpha \wedge \beta^i_j$ to exist (for at least one j by the definition of injectivity) and so this forces β^i_j to hold in A.

6.5.5 Proof

Theorem 6.5.5.1. Let $\phi(x_1, \ldots, x_d)$ be a first-order formula of the form

$$\bigwedge_{1 \leq i \leq z} \left(\alpha_i \rightarrow \bigvee_{1 \leq j \leq p_i} \beta^i_j \right),$$

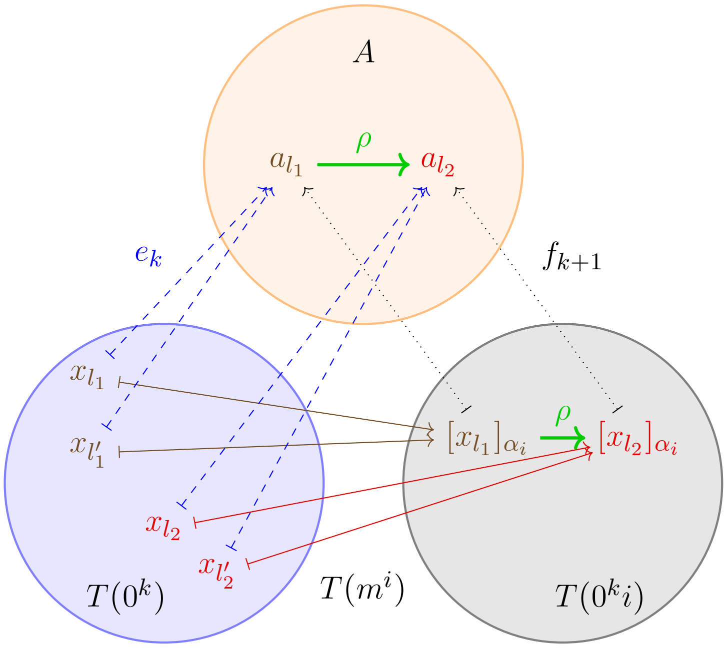
where $z \in \mathbb{N}_0$ and each $p_i \in \mathbb{N}_0$, such that each α_i is a conjunction of atomic formulae, and each β^i_j is an atomic formula. Further, let $n \in \mathbb{N}^+$ be such that $d < 2n$ and define $k = 2n$.

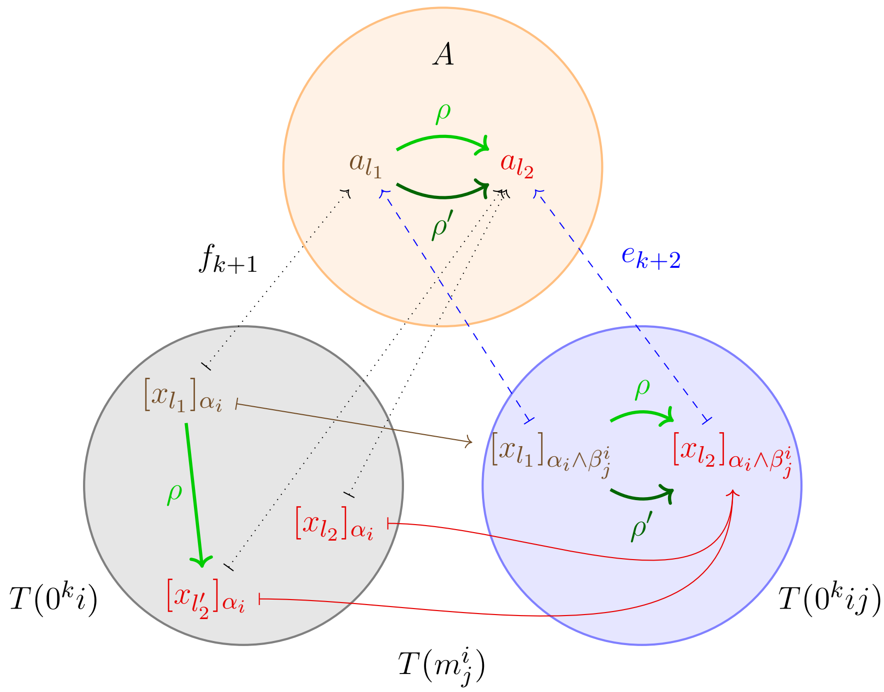
Define the tree T_ϕ with image in $\mathbb{C} = \mathbf{RStr}$ as above.

Then, for any $A \in \mathbb{C}$,

$$\forall_{x_1}\exists_{x_2}\cdots,\forall_{x_{k-1}}\exists_{x_k}[\phi(x_1,\ldots,x_d)] \qquad \text{iff} \qquad A \models T_\phi.$$

Proof. It may be useful to refer to the following three diagrams:





$$\begin{tikzcd}
 & & & A & \\
T(0) \arrow[r, "T(\bar{m}^1)"] \arrow[urrr, bend left, "\forall f_1"] & \cdots \arrow[r, "T(\bar{m}^k)"] & T(0^k) \arrow[r, "T(m^i)"] \arrow[ur, "\exists e_k"] & T(0^k i) \arrow[r, "T(m^i_j)"] \arrow[u, "\forall f_{k+1}"] & T(0^k ij) \arrow[ul, bend right, "\exists e_{k+2}"']
\end{tikzcd}$$

($\Rightarrow$) Assume $\forall_{x_1}\exists_{x_2}\cdots,\forall_{x_{k-1}}\exists_{x_k}[\phi(x_1,\ldots,x_d)]$ holds for A.

Let ρ denote the n-ary (for some $n \in \mathbb{N}$) relation corresponding to the interpretations, in $T(0^k i)$, $T(0^k ij)$ and A of n-ary symbol ρ.

Then for each $f_1 : T(0) \to A$, assume $f_1(x_1^1) = a_1$, for a specific element $a_1 \in A$, for the unique $x_1^1 \in T(0)$. This is trivially a morphism since $T(0)$ has no relations.

For any $a_2 \in A$, there is a map $e_2 : T(0^2) \to A$ with $e_2(x_1^2) = a_1 = f_1(x_1^1)$ and we may assume $e_2(x_2^2) = a_2$. In particular, let a_2 be an element of A (depending on a_1) asserted by the existence quantifier in the fragment $\forall_{x_1}\exists_{x_2}$. (Which exists by the truth of the sentence given in the assumption).

These are the only two elements in $T(0^2)$. Again, e_2 is trivially a morphism. It commutes ($e_2 T(\bar{m}^2) = f_1$) by construction.

If $k > 2$, then for any $f_3 : T(0^3) \to A$ with $f_3 T(\bar{m}^3) = e_2$, there is, for any $a_4 \in A$, a morphism $e_4 : T(0^4) \to A$ for which commutativity holds, and for which $e_4(x_4^4) = a_4$. Again, let a_4 be an element given by $\exists_{x_4}$ in $\forall_{x_1}\exists_{x_2}\forall_{x_3}\exists_{x_4}$.

This process continues up to the morphism $e_k : T(0^k) \to A$.

Let $a_1,\ldots,a_k \in A$ be the specific tuple of elements such that for every $1 \le i \le k$, it is true that $e_k(x_i^k) = a_i$. Then these elements have been chosen such that $\phi(a_1,\ldots,a_d)$ holds.

Then, since $\phi(a_1,\ldots,a_d)$ holds, the conjunction

$$\bigwedge_{i \le z}\left(\alpha_i \to \bigvee_{j \le p_i} \beta_j^i\right)$$

holds for these elements as well.

Let $1 \le i \le z$, and let $f_{k+1} : T(0^k i) \to A$ be any morphism which commutes as $f_{k+1}T(m^i) = e_k$.

Finally, it is shown that there are $1 \le j \le p_i$ and morphism $e_{k+2} : T(0^k ij) \to A$ which commutes as $e_{k+2}T(m_j^i) = f_{k+1}$.

Since f_{k+1} is a morphism, then by definition of $T(0^k i)$, this implies that α_i holds for $a_1,\ldots,a_d$. Hence, then also

$$\bigvee_{j \le p_i} \beta_j^i$$

holds, and so for a particular $j \leq p_i$, the fragment β^i_j holds. Fix this j.

An element of $T(0^k i)$ can be thought of as an equivalence class $[x^k]_{\alpha_i}$ and an element of $T(0^k ij)$ as an equivalence class $[x^k]_{\alpha_i \wedge \beta^i_j}$ or an equivalence class $[[x^k]_{\alpha_i}]_{\beta^i_j}$.

The morphism $T(m^i_j)$ takes two elements $[x^k]_{\alpha_i}$ and $[(x^k)']_{\alpha_i}$ to the same element by definition precisely if for some x^k in the first equivalence class and $(x^k)'$ in the second there are corresponding $a = e_k(x^k)$ and $a' = e_k((x^k)')$ asserted to be equal by β^i_j.

If $T(m^i_j)$ maps $[x^k]_{\alpha_i}$ and $[(x^k)']_{\alpha_i}$ in $T(0^k i)$ to the same element in $T(0^k ij)$, then x^k and $(x^k)'$ are mapped to the same element by e_k, and hence by commutativity, then f_{k+1} also maps $[x^k]_{\alpha_i}$ and $[(x^k)']_{\alpha_i}$ to the same element in A.

Hence it is possible to define a map $e_{k+2} : T(0^k ij) \to A$, that commutes by $e_{k+2}T(m^i_j) = f_{k+1}$, on each element as

$$e_{k+2}([x^k]_{\alpha_i \wedge \beta^i_j}) = f_{k+1}([x^k]_{\alpha_i})$$

and this definition is well-defined by the above argument.

Assume, for relation symbol ρ which is a subformula of α_i, that the elements $[x^k_{l_1}]_{\alpha_i \wedge \beta^i_j}, \dots, [x^k_{l_n}]_{\alpha_i \wedge \beta^i_j}$ of $T(0^m ij)$ are related by ρ. Then there are representatives $[x^k_{l_1}]_{\alpha_i}, \dots, [x^k_{l_n}]_{\alpha_i}$ which are related by ρ in $T(0^m i)$.

Since f_{k+1} is a morphism, then the images $a_{l_1} = f_{k+1}([x^k_{l_1}]_{\alpha_i})$, $\dots$, $a_{l_n} = f_{k+1}([x^k_{l_n}]_{\alpha_i})$ are related by ρ in A.

Hence e_{k+2} is a morphism.

($\Leftarrow$) Assume $A \models T_\phi$.

Let $f_1 : T(0) \to A$ be a morphism taking the unique element x^1_1 to some $a_1 \in A$. There is some $a_2 \in A$ given by where the morphism $e_2 : T(0^2) \to A$ takes x^2_2 (it must take x^2_1 to the same place as f_1 takes x^1_1). Continuing in this fashion, obtain $a_1, \dots, a_k \in A$ corresponding to the subscripts of the quantifiers $\forall_{x_1} \exists_{x_2} \cdots \forall_{x_{k-1}} \exists_{x_k}$.

Now, fix i and assume that α_i is true of $a_1, \dots, a_k$. Let f_{k+1} be a morphism $T(0^k i) \to A$ commuting in the right way, and whose existence follows from the truth of α_i.

Then there is j and morphism $e_{k+2} : T(0^k ij) \to A$ commuting in the right way.

Now it is shown that β^i_j holds for $a_1, \dots, a_k$.

If β^i_j states that $a_{l_1} = a_{l_2}$ for some l_1 and l_2, then $T(m^i_j)T(m^i)$ maps x_{l_1} and x_{l_2} to the same image $[x]_{\alpha_i \wedge \beta^i_j}$ in $T(0^k ij)$. Since e_{k+2} can map this image to only one element in A, then by commutativity, the morphism $e_k : T(0^k) \to A$ must map them to the same place, and so $a_{l_1} = a_{l_2}$.

If ρ is an n-ary relation which is a subformula of β^i_j, and it asserts that the variables $x_{l_1}, \dots, x_{l_n}$ are related, then the elements $[x_{l_1}]_{\alpha_i \wedge \beta^i_j}, \dots, [x_{l_n}]_{\alpha_i \wedge \beta^i_j}$ are related by ρ in $T(0^k ij)$ and so the images of these elements under e_{k+2} must be related by ρ in A. But by definition of relations in $T(0^k ij)$, this means there are representatives $x_{l_1}, \dots, x_{l_n}$ whose images $a_{l_1}, \dots, a_{l_n}$ under e_k must related by ρ in A.

Hence $\alpha_i \to \beta^i_j$ holds in A and so

$$\alpha_i \to \bigvee_{j \le p_i} \beta^i_j$$

does too. But this argument worked for all i, and so the conjunction

$$\bigwedge_{i \le z} \left(\alpha_i \to \bigvee_{j \le p_i} \beta^i_j \right)$$

holds for these elements as well. □

Remark. The above theorem and proof work even for nullary relations. Even if ϕ is a sentence consisting of only nullary relations, it is possible to simply choose variables x_1 and x_2 not named in the formula and append $\forall_{x_1} \exists_{x_2}$ in from of ϕ. This works provided that A is nonempty.

6.5.6 Empty Structures

All of the above works for A nonempty. Some minor adjustments can be made to the above definitions and theorems to take into account empty A.

The following is an informal description of the changes required to effect this.

First note that no sentence beginning with an existential quantifier will be true of an empty structure, whereas any sentence beginning with a universal quantifier will be.

For non-empty A, if we wanted to capture a sentence without quantifiers or a sentence beginning with an existence quantifier, we could just add quantifiers to the front which refer to variables not mentioned in ϕ. But by the previous paragraph, this may change the formula for empty A.

To produce a tree from a formula beginning with an existential quantifier, we add an empty relational structure $\mathcal{E}$ to the beginning of the image of the tree we might otherwise form. There is always a unique morphism from this empty structure to each of A and $T(0)$. Furthermore, these morphisms will commute with any $f_1 : T(0) \to A$ (which should actually now be labelled e_1 since it will correspond to an existential quantifier). Now, if we want to form a tree for a formula $\exists_{x_1}\forall_{x_2}\cdots\phi$, add an imaginary $\forall$ on the beginning and form the tree as usual, with an empty structure adjoined at the beginning corresponding to an imaginary $\forall$.

In this extension, the trees so formed from the sentences beginning with an existential quantifier will work for arbitrary A, empty or not.

$$\begin{tikzcd}
 & & & & A & \\
\mathcal{E} \arrow[r, dotted, "\text{empty}"] \arrow[urrrr, dotted, bend left, "\forall f_0"] & T(0) \arrow[r, "T(\bar{m}^1)"] \arrow[urrr, bend left, "\exists e_1"] & \cdots \arrow[r, "T(\bar{m}^k)"] & T(0^k) \arrow[r, "T(m^i)"] \arrow[ur, "\forall f_k"] & T(0^k i) \arrow[r, "T(m^i_j)"] \arrow[u, "\exists e_{k+1}"'] & T(0^k ij) \arrow[ul, bend right, "\forall f_{k+2}"']
\end{tikzcd}$$

What now, if a sentence without quantifiers is to be converted to a tree? Then the sentence may be true of certain empty relational structures A but not others (depending on whether each nullary relation ρ in the language is true of A or not).

In this case, it suffices to remove all of the objects in the image of the tree corresponding to quantifiers, and to add two empty structures to the beginning.

$$\begin{tikzcd}
 & & A & \\
\mathcal{E}_1 \arrow[r, dotted, "\text{empty}"] \arrow[urr, dotted, bend left, "\forall f_0"] & \mathcal{E}_2 \arrow[r, dotted, "\text{empty}"] \arrow[ur, dotted, bend left, "\exists e_0"] & T(i) \arrow[r, "T(m^i_j)"] \arrow[u, "\forall f_1"'] & T(ij) \arrow[ul, bend right, "\exists e_2"']
\end{tikzcd}$$

However, Definition 6.5.3.3 does not allow for empty structures in the image of the tree. Hence, define three cases depending on whether the sentence to be converted to a tree has zero, one, or more than one quantifier. (The definition of injectivity, however, does not change)

For more than one quantifier, the definition is exactly as before.

For exactly one quantifier, add an extra object (also denoted $\mathcal{E}_1$) to the beginning of the tree category $\mathbb{T}_\phi$ that would otherwise have been formed (in the definition of a tree from §6.5.3). Now define $T(\mathcal{E}_1) = \langle \emptyset, (R^\emptyset)^\emptyset \rangle$, and $T(\text{empty})$ = empty. the other objects in the tree category are mapped as before.

For zero quantifiers (when ϕ is a sentence), add two objects $\mathcal{E}_1$ and $\mathcal{E}_2$ to the beginning of the tree category, then do the same as in the previous case.

$$\begin{tikzcd}
 & & & & A & \\
T(\mathcal{E}_1) \arrow[r, dotted, "\text{empty}"] \arrow[urrrr, dotted, bend left, "\forall f_0"] & T(\mathcal{E}_2) \arrow[r, dotted, "\text{empty}"] \arrow[urrr, dotted, bend left, "\exists e_0"] & \cdots \arrow[r, "T(\bar{m}^k)"] & T(0^k) \arrow[r, "T(m^i)"] \arrow[ur, bend left, "\exists e_k"] & T(0^k i) \arrow[r, "T(m^i_j)"] \arrow[u, "\forall f_{k+1}"] & T(0^k ij) \arrow[ul, bend right, "\exists e_{k+2}"']
\end{tikzcd}$$

Hence, allowing k to be a minimum of 0 instead of a minimum of 2, and placing either one or two $\mathcal{E}s$ at the front of the tree, this allows an arbitrary formula to be converted to a tree which will work for all structures, empty or not.

Chapter 7

Łoś's Theorem in Every Category

7.1 Overview

The following chapter is primarily an exposition of the paper [AN78]. The primary theorems of that paper are stated and their proofs are given in greater detail. New diagrams are given which do not appear in the paper, as well as intuition, both of which are intended to clarify the proofs.

Additionally, the results of that paper are applied to give an entirely category-theoretic proof that the (category-theoretic) ultraproduct of fields is a field. This proof serves also as a demonstration of how trees (from [AN79]) can be used to distinguish objects with certain properties in a category, via the existence of certain morphisms.

In this case, trees are used to distinguish fields in the category of rings, and the version of Łoś's Theorem from [AN78] is then applied to show that every ultraproduct of fields (in the category of rings) is itself a field. The author does not know of any source outside of this dissertation which contains this specific example.

The main theorem of this chapter is Theorem 7.3.1.1. It restates Łoś's Theorem in terms of injectivity with respect to trees instead of truth of formulae. It says that an ultraproduct is injective with respect to a given (compact) tree if and only if a 'large' (ultrafilter-indexed) family of objects from the ultraproduct are injective with respect to the same tree.

This theorem only holds for a special class of trees, known as 'compact' trees (in [AN79], referred to as 'strongly small' trees). Such a tree captures the intuition of 'finitely presentable' objects. In a category of relational structures, it corresponds to finite (and empty) relational structures, and trees comprising such structures correspond exactly to finitary first-order formulae (formulae which can be written using finitely many symbols).

The main theorem and its proof involve that it is possible to factor any "assignment" morphism from a compact tree into the ultraproduct through an "assignment" morphism into some large-indexed product, and out again through a coprojection. Such a morphism into a product can then be projected onto a large family of the objects from which the ultraproduct is formed.

The proof goes by first proving a version of Łoś's Theorem involving specific "assignment" morphisms. This corresponds to the fact that a first-order formula $\phi(x_1, \dots, x_n)$ in n variables holds for an assignment $\phi([a_1], \dots [a_n])$ of equivalence classes $[a_k] \in \prod_{\mathcal{U}} A_i$ in the ultraproduct if and only if, for some large family $(A_i)_{i \in Z}$ (indexed by $Z \in \mathcal{U}$), it holds for the assignments $\phi(a_1, \dots, a_n)$ of representatives $a_1, \dots, a_n \in A_i$ of the equivalence classes $a_k \in [a_k]$.

From the fact that this works for any such assignment, the "sentence" version of Łoś's theorem then follows.

7.2 Compactness

7.2.1 Definitions

The original paper [AN79], requires that elements of the tree be **strongly small**. (The reader can look in [HS73] for more information on this property).

The definition of "strongly-small" uses the concept of a directed category. The definition of "compact" is similar, but uses the more modern concept of a filtered category. By the paper [AN82], the two definitions are equivalent.

Other terms for compactness which are used in the literature include **finitely presented**, **finitely presentable** and **of finite presentation**.

To define compactness, we first need the concept of a filtered category.

Definition 7.2.1.1. A category $\mathbb{I}$ is a **filtered category** if it is a non-empty category (a category with at least one object) such that there is an upper bound for every pair of objects and every pair of parallel morphisms.

In other words: for every pair of objects $I, J \in \mathbb{I}$ there exists an object $K \in \mathbb{I}$ and morphisms $i : I \to K$ and $j : J \to K$.

For every pair objects $I, J \in \mathbb{I}$ and pair of parallel morphisms $f, g : I \to J$ there is an object K and morphism $d : J \to K$ such that $df = dg$. $\triangleleft$

$$\begin{tikzcd}
 & \vdots & \\
 & K \arrow[u] & \\
 & \vdots \arrow[u] & \\
I \arrow[uur, bend left, "i"] & & J \arrow[uul, bend right, "j"']
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
I \arrow[r, bend left, "f"] \arrow[r, bend right, "g"'] & J \arrow[r, "d"] & K
\end{tikzcd}$$

Remark. A filtered category generalizes the notion of a directed poset. ◁

Remark. Being filtered is weaker than requiring that coproducts and coequalizers exist. There may exist some K with $i : I \to K$ and $j : J \to K$ and yet not exist universal such K, i, j. Similarly, there may exist K and d such that $df = dg$ but no universal such K and d. ◁

A routine argument shows the following characterization (which is often taken to be the definition of a filtered category) is equivalent.

Proposition 7.2.1.2. A category is filtered if and only if every finite diagram has a cone out of that diagram.

In other words, for any diagram $F : \mathbb{D} \to \mathbb{I}$ there is an object K and family of morphisms $\langle k_D : F(D) \to K\rangle_{D\in\mathbb{D}}$, and each morphism k_D commutes with the diagram in the sense that, for any pair of objects $C, D \in \mathbb{D}$ and morphism $f : C \to D$, it holds that $k_D\ F(f) = k_C$. ◁

Definition 7.2.1.3. A colimit $\varinjlim D$ of a diagram $D : \mathbb{I} \to \mathbb{C}$, is a **filtered colimit** if the domain $\mathbb{I}$ is a filtered category. ◁

Definition 7.2.1.4. Given a category $\mathbb{C}$, an object $A \in \mathbb{C}$ is a **compact object** if, for any filtered colimit $E = \varinjlim D$ of a diagram $D : \mathbb{I} \to \mathbb{C}$ where

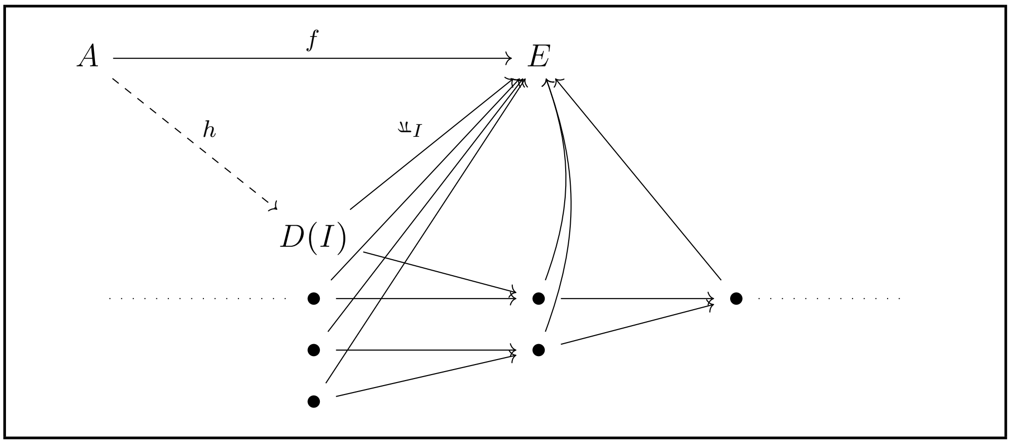
$$\langle \pitchfork_I : D(I) \to E\rangle_{I\in\mathbb{I}}$$

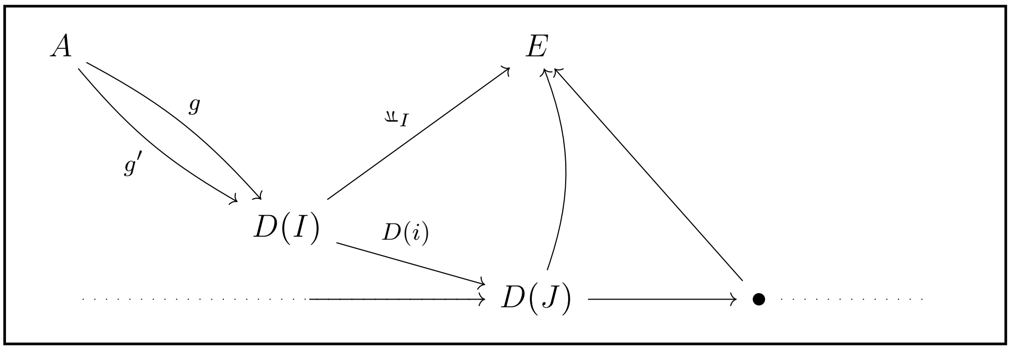
is the colimiting cone, the following two conditions are satisfied:

1. Every arrow $f : A \to E$ factors through the cone. I.e., there are $I \in \mathbb{I}$ and morphism $h : A \to D(I)$ (with $\pitchfork_I : D(I) \to E$) such that $f = \pitchfork_I h$.

2. For any $I \in \mathbb{I}$ and any pair of morphisms $g, g' : A \to D(I)$, such that $\pitchfork_I g = \pitchfork_I g'$, there is some $J \in \mathbb{I}$ and morphism $i : I \to J$ such that $D(i)\ g = D(i)\ g'$.

◁





Definition 7.2.1.5. A tree $T : \mathbf{T} \to \mathbb{C}$ is **compact** if every object $T(I)$ in the image of the tree is compact. ◁

Remark. Compactness captures the intuition of an object being determined by a finite amount of information. This is important, because we want the trees comprising such objects to correspond to (or generalize) finitary formulae.

The word compact is used because of the resemblance to compactness in topology. If a compact topological space is covered by a collection of open sets, then it is covered by a finite subcollection (and, if this collection of open sets is closed under finite unions, then the compact space can be covered by a single open set). Similarly, if morphism $f : A \to E$ is seen as a 'covering' of A by the family of objects in the directed diagram, along with their morphisms, then compactness says that it can be covered by finitely many of them. In fact, it is possible to cover it with just one.

From an algebraic viewpoint, the definition of compactness captures the concept of finitely presented. Suppose that A is a finitely presented algebraic object (such as a finitely presented group), and suppose that E was an algebraic object which is not finitely presented. Then E can be written as a directed colimit (direct limit in algebraic terms) of all its sub-objects (e.g.

subgroups). In particular, suppose it could be written as a directed colimit of subobjects, all of which are strictly smaller. Then any morphism $f : A \to E$ factors through this cone in a nice way, so the information of f is already determined by strictly less information than is contained in E. $\triangleleft$

Remark. The above definition is similar to the one given in 7.1 for **strongly small** objects, though that paper uses a condition involving directed colimits rather than filtered colimits. By the results of the paper [AN82], the two conditions are equivalent.

The version in terms of filtered colimits is the more natural one, category-theoretically speaking. $\triangleleft$

7.2.2 Proof of Equivalent Definition of Compactness

The following is the main theorem of this subsection.

This theorem is not necessary for the proof of Łoś's Theorem, and so its proof may be safely skipped. However, it provides the more commonly used definition of compactness, and shows that it matches the one given above.

Theorem 7.2.2.1. For an object $C \in \mathbb{C}$, the following are equivalent:

1. C is compact
2. The functor $\mathrm{Hom}(C, -)$ commutes with filtered colimits.

$\triangleleft$

The above theorem makes use of the following definition:

Definition 7.2.2.2. A functor

$$F : \mathbb{C} \to \mathbb{D}$$

is said to **preserve filtered colimits**, or to **commute with filtered colimits** if, for every diagram $D : \mathbb{I} \to \mathbb{C}$ from a filtered category $\mathbb{I}$, it holds that

$$\varinjlim FD \quad \cong \quad F(\varinjlim D)$$

and the corresponding cones are isomorphic.

What this means is that the morphisms for the colimiting cone are given by the cone of images $\langle F(\rotatebox[origin=c]{180}{π}_I) : FD(I) \to F(\varinjlim D)\rangle_{I \in \mathbb{I}}$ under F of the colimiting cone $\langle \rotatebox[origin=c]{180}{π}_I : D(I) \to \varinjlim D\rangle_{I \in \mathbb{I}}$.

Here $\rotatebox[origin=c]{180}{π}_I$ is the coprojection of the colimiting cone. (The notation is not entirely standard). $\triangleleft$

Remark. The definition of a functor commuting with directed colimits is the same as the above definition where **filtered** is replaced with **directed**. $\triangleleft$

The following proofs were written by the author of this dissertation, though it is possible a similar proof may exist in [Mat87] for the case of directed colimits, since that paper is referenced by [AN78], as proof of the statement made in Definition 5 of the latter paper (definition of strongly small objects).

Proposition 7.2.2.3. If an object $C \in \mathbb{C}$ is compact, then the functor

$$\mathrm{Hom}(C, -) : \mathbb{C} \to \mathbf{Set}$$

preserves filtered colimits.

In other words, for every diagram $D : \mathbb{I} \to \mathbb{C}$ (from a filtered category $\mathbb{I}$) it holds that

$$\varinjlim \mathrm{Hom}(C, D(-)) \qquad \cong \qquad \mathrm{Hom}(C, \varinjlim D)$$

and the coprojections for the colimit are given as

$$\mathrm{Hom}(C, \amalg_I) : \quad \mathrm{Hom}(C, D(I)) \qquad \to \qquad \mathrm{Hom}(C, \varinjlim D).$$

Proof. Let $D : \mathbb{I} \to \mathbb{C}$ be a filtered colimit.

Assume $C \in \mathbb{C}$ is compact.

First, the fact that the family of images of the coprojections of the colimiting cone itself forms a cone follows simply from the fact that $\mathrm{Hom}(C, -)$ is a functor and hence preserves composition:

Let $f : I \to J \in \mathbb{C}$. Then $\amalg_J\, D(f) = \amalg_I$ by definition, and so

$$\mathrm{Hom}(C, \amalg_J)\mathrm{Hom}(C, D(f)) \quad = \quad \mathrm{Hom}(C,\ \amalg_J D(f)) \quad = \quad \mathrm{Hom}(C, \amalg_I).$$

Now, in order to show that $\mathrm{Hom}(C, \varinjlim D)$ is the colimit in **Set**, it is necessary to define, for each object $W \in \mathbf{Set}$ with corresponding cone, a map

$$\gamma : \mathrm{Hom}(C, \varinjlim D) \to W$$

which commutes in the correct way, and it must be shown this map is the unique map doing such.

$$\begin{tikzcd}
 & W & \\
 & \mathrm{Hom}(C, \varinjlim D) \arrow[u, dashed, "\gamma"] & \\
\mathrm{Hom}(C, D(I)) \arrow[ur, "\mathrm{hom}(C,\amalg_I)"] \arrow[uur, bend left, "\omega_I"] \arrow[rr, "\mathrm{Hom}(C,D(f))"'] & & \mathrm{Hom}(C, D(J)) \arrow[ul, "\mathrm{hom}(C,\amalg_J)"'] \arrow[uul, bend right, "\omega_J"']
\end{tikzcd}$$

Let $W \in \mathbf{Set}$ with cone $\langle \omega_I : \mathrm{Hom}(C, D(I)) \to W \rangle_{I \in \mathbb{I}}$.

By compactness, for each $f : C \to \varinjlim D$ there exist an $I \in \mathbb{I}$ and $h : C \to D(I)$ such that $f = \amalg_I h$.

Now it is shown that $\omega_I(h)$ is independent of the specific I and h chosen (such that $f = \amalg_I h$).

Let I' and $h' : C \to D(I')$ be such that $f = \amalg_{I'} h'$. Then, since $\mathbb{I}$ is filtered, there are object K and morphisms $i : I \to K$ and $i' : I' \to K$. Now, $\amalg_K D(i) = \amalg_I$ and $\amalg_K D(i') = \amalg_{I'}$ by definition of $\amalg$ being a cone.

Now,

$$\amalg_K D(i)h = \amalg_I h = f = \amalg_{I'} h' = \amalg_K D(i')h'$$

Since $\amalg_K(D(i)h) = \amalg_K(D(i')h')$, then by the second property of compactness there are object J and morphism $k : K \to J$ such that $D(k)D(i)h = D(k)D(i')h'$.

$$\begin{tikzcd}
C \arrow[rr, "f"] \arrow[dr, "h"] \arrow[dddr, "h'"'] & & \varinjlim D & & \\
& D(I) \arrow[ur, "\amalg_I"] \arrow[drr, "D(i)"] & & & \\
& & & D(K) \arrow[uul, "\amalg_K"] \arrow[r, "D(k)"] & D(J) \arrow[uull, "\amalg_J"'] \\
& D(I') \arrow[uuur, "\amalg_{I'}"] \arrow[urr, "D(i')"'] & & &
\end{tikzcd}$$

Hence, also the equality

$$\mathrm{Hom}(C, D(k))\,\mathrm{Hom}(C, D(i))(h) = \mathrm{Hom}(C, D(k))\,\mathrm{Hom}(C, D(i'))(h')$$

holds.

Since ω is a cone, it holds that

$$\omega_I = \omega_J \mathrm{Hom}(C, D(k)) \mathrm{Hom}(C, D(i))$$

and

$$\omega_{I'} = \omega_J \mathrm{Hom}(C, D(k)) \mathrm{Hom}(C, D(i'))$$

Hence, $\omega_I(h) = \omega_{I'}(h')$.

Define the map

$$\gamma : \mathrm{Hom}(C, \varinjlim D) \to W$$

to map each $f : C \to \varinjlim D$ to this unique $w \in W$ where $w = \omega_I(h)$ for some I and h such that $f = \text{ш}_I h$.

Then, for any I, of course $\omega_I = \gamma\mathrm{Hom}(C, \text{ш}_I)$, since for any $h : C \to D(I)$, it holds that $\gamma\mathrm{Hom}(C, \text{ш}_I)(h) = \gamma(\text{ш}_I h) = \omega_I(h)$ by definition of γ.

That this map γ must be unique also follows directly from the above. For any $f : C \to \varinjlim D$ there are h and I for which $f = \text{ш}_I h$. If, for all I it holds that $\gamma\mathrm{Hom}(C, \text{ш}_I) = \omega_I$, then, in particular, for all $h : C \to D(I)$ also $\gamma\mathrm{Hom}(C, \text{ш}_I)(h) = \omega_I(h)$. So, for each f there are I and h, such that $\gamma(f) = \gamma(\text{ш}_I h) = \gamma\mathrm{Hom}(C, \text{ш}_I)(h) = \omega_I(h)$. $\square$

Remark. The above proof can be used, with almost no adaptation, to prove that the object C being strongly small (see §7.2.1) implies that the Hom functor preserves colimits. $\lhd$

For the next proposition, we shall first need two lemmas.

Lemma 7.2.2.4. In the category **Set**, a colimit of a diagram $D : \mathbb{I} \to \mathbf{Set}$ is given by:

$$\left(\coprod_{I \in \mathbb{I}} D(I)\right) \Big/ \sim$$

where $\coprod$ is the disjoint union and $\sim$ is the equivalence relation generated by

$$\{\langle x, y\rangle \mid I, J \in \mathbb{I}\ ;\ x \in D(I); y \in D(J);$$
$$\exists K(\exists i : I \to K[\exists j : J \to K(D(i)(x) = D(j)(y))]);\}.$$

In particular, if D is a filtered diagram, then the above set is an equivalence relation and so $\sim$ is **equal** to that set.

Lemma 7.2.2.5. Let E be a colimit of a filtered diagram $D : \mathbb{I} \to \mathbf{Set}$. Let $x, x' \in D(I)$ be two elements such that $\text{ш}_I(x) = \text{ш}_I(x')$.

Then there is some $J \in \mathbb{I}$ and morphism $i : I \to J$ such that $D(i)(x) = D(i)(x')$.

Proof. Since, by Lemma 7.2.2.4, the colimit is given by the disjoint union quotiented by the equivalence relation

$$\{\langle x, y\rangle \mid I, J \in \mathbb{I}\ ;\ x \in D(I); y \in D(J);$$
$$\exists K(\exists i : I \to K[\exists j : J \to K(D(i)(x) = D(j)(y))]);\},$$

then, in particular, the pair (x, x') must appear in this equivalence relation. Because it is a directed colimit, the equivalence relation generated by the above set is equal to that set. In other words, there is J and $i : I \to J$ such that $D(i)(x) = D(i)(x')$. $\square$

Proposition 7.2.2.6. If the functor $\mathrm{Hom}(C, -)$ preserves filtered colimits, then C is compact.

Proof. Let $D : \mathbb{I} \to \mathbb{C}$ be a filtered diagram.

In **Set**, the cone for a colimit is jointly surjective onto the colimit, since, by Lemma 7.2.2.4, the colimit is a quotient of the disjoint union of objects in the diagram. Hence, each $f \in \mathrm{Hom}(C, \varinjlim D)$ is, for some $I \in \mathbb{I}$, the image under $\mathrm{Hom}(C, \amalg_I)$ of some $h \in \mathrm{Hom}(C, D(I))$.

Thus, for each $f : C \to \varinjlim D$, there are $I \in \mathbb{I}$ and $h : C \to D(I)$ such that $f = \amalg_I h$.

Now note that $\mathrm{Hom}(C, D(-))$ is a functor $\mathbb{I} \to$ **Set**. Hence that diagram is also filtered, so the first hypothesis of 7.2.2.5, holds.

Now, let $g, g' \in \mathrm{Hom}(C, D(I))$ such that $\amalg_I g = \amalg_I g'$.

Then by Lemma 7.2.2.5, there must be some $i : I \to J$ such that

$$\mathrm{Hom}(C, D(i))(g) = \mathrm{Hom}(C, D(i))(g'),$$

and so $D(i)g = D(i)g'$. $\square$

Remark. Again, the above proof can be adapted almost without change to show that if the Hom functor preserves directed colimits, then C is strongly small (see §7.2.1). $\triangleleft$

The proof of Theorem 7.2.2.1 then follows directly from the above.

Proof of Theorem 7.2.2.1.

- (1) $\Rightarrow$ (2) is given by Proposition 7.2.2.3.
- (2) $\Rightarrow$ (1) is given by Proposition 7.2.2.6.

$\square$

Lemma 7.2.2.7. Viewed as a category, an ultrafilter is (downward) filtered.

Proof. Let $X, Y \in \mathcal{U}$. Then $Z = X \cap Y \in \mathcal{U}$. Also, the morphisms are unique, hence it is directed and so it is filtered. $\square$

7.2.3 Lemmas For Compact Objects and Ultraproducts

The following is a very useful lemma regarding maps from compact objects into the diagram of which a given ultraproduct is the colimit. This fact is used implicitly (without comment) in the proofs of the paper [AN78], but is stated explicitly as a lemma here.

Intuitively, what it says is that if two families $\langle f_i \rangle_{i\in X}$ and $\langle g_i \rangle_{i\in Y}$ out of a compact object C are "essentially the same" according to the ultraproduct, then they agree on some large set Z.

$$\begin{tikzcd}
 & & \prod_{\mathcal{U}} A_i & & \\
 & \prod_X A_i \arrow[ur, "\underline{\psi}_X"] \arrow[drr, "\pi^X_W"] & & & \\
C \arrow[ur, "\langle f_i \rangle_{i\in X}"] \arrow[r, "\langle g_i \rangle_{i\in Y}"] \arrow[rrrr, bend right, "\langle f_i \rangle_{i\in Z} = \langle g_i \rangle_{i\in Z}"'] & \prod_Y A_i \arrow[uur, "\underline{\psi}_Y"] \arrow[rr, "\pi^Y_W"] & & \prod_W A_i \arrow[uul, "\underline{\psi}_W"'] \arrow[r, "\pi^W_Z"] & \prod_Z A_i \arrow[uull, "\underline{\psi}_Z"']
\end{tikzcd}$$

Lemma 7.2.3.1. Let $C \in \mathbb{C}$ be a compact object, let $(A_i)_{i\in I}$ be a family of objects in $\mathbb{C}$, and let $\prod_{\mathcal{U}} A_i$ be an ultraproduct formed from this family with some ultrafilter $\mathcal{U} \subseteq \mathcal{P}(I)$.

Let $X, Y \in \mathcal{U}$, and let

$$\langle f_i \rangle_{i\in X} : \qquad C \to \prod_X A_i$$

$$\langle g_i \rangle_{i\in Y} : \qquad C \to \prod_Y A_i$$

be two morphisms such that

$$\underline{\psi}_X \langle f_i \rangle_{i\in X} \quad = \quad \underline{\psi}_Y \langle g_i \rangle_{i\in Y}.$$

Then, there is $Z \in \mathcal{U}$ (with $Z \subseteq X \cap Y$) such that

$$\langle f_i \rangle_{i\in Z} \quad = \quad \langle g_i \rangle_{i\in Z},$$

i.e., such that

$$f_i = g_i, \qquad \text{for all } i \in Z.$$

Proof. First, note that the diagram over which an ultraproduct is taken is upward filtered (since it is a contravariant functor from an ultrafilter, which is downwards filtered).

Let $W = X \cap Y$, then

$$\begin{aligned}\Perp_W \, \pi^X_W \, \langle f_i \rangle_{i \in X} &= \Perp_X \, \langle f_i \rangle_{i \in X} \\ &= \Perp_Y \, \langle g_i \rangle_{i \in Y} \\ &= \Perp_W \, \pi^Y_W \, \langle g_i \rangle_{i \in Y}\end{aligned}$$

Now, by the second property of compactness of C, applied to morphisms $\pi^X_W \, \langle f_i \rangle_{i \in X}$ and $\pi^Y_W \, \langle g_i \rangle_{i \in Y}$ there exists Z such that

$$\pi^W_Z \, \pi^X_W \, \langle f_i \rangle_{i \in X} \quad = \quad \pi^W_Z \pi^Y_W \, \langle g_i \rangle_{i \in Y}.$$

Hence,

$$\begin{aligned}\langle f_i \rangle_{i \in Z} &= \pi^X_Z \, \langle f_i \rangle_{i \in X} \\ &= \pi^W_Z \, \pi^X_W \, \langle f_i \rangle_{i \in X} \\ &= \pi^W_Z \pi^Y_W \, \langle g_i \rangle_{i \in Y} \\ &= \pi^Y_Z \, \langle g_i \rangle_{i \in Y} \\ &= \langle g_i \rangle_{i \in Z}\end{aligned}$$

□

One of the most useful cases of this lemma is via the following corollary. Basically, the corollary says that if a morphism c into the ultraproduct from a compact object C factors as bm via some other compact object B, then it is possible to find some large family $\langle A_i \rangle_{i \in Z}$ of objects (of the ultraproduct) indexed by Z and two families $\langle c_i : C \to A_i \rangle_{i \in Z}$ and $\langle b_i : B \to A_i \rangle_{i \in Z}$ of morphisms from C into the family of objects, which have certain properties.

$$
\begin{tikzcd}
 & \prod_{\mathcal{U}} A_i & \\
C \arrow[ur, "c"] \arrow[rr, bend left, "\langle c_i\rangle_{i\in Z}"] \arrow[rr, bend right, "\langle b_i m\rangle_{i\in Z}"'] \arrow[d, "m"'] \arrow[ddrr, bend left, "c_i"] \arrow[ddrr, bend right, "b_i m"'] & & \prod_{Z} A_i \arrow[ul, "\perp\!\!\!\perp_Z"'] \arrow[dd, "\pi_i"] \\
B \arrow[uur, "b"] \arrow[urr, "\langle b_i\rangle_{i\in Z}"'] \arrow[drr, "b_i"'] & & \\
 & & A_i
\end{tikzcd}
$$

These families of morphisms have the properties that firstly, they contain "all the information of c and b" as far as the ultraproduct is concerned, and secondly, $c_i = b_i m$ for each $i \in Z$.

This becomes useful specifically in the case where $C = T(X_1)$ and $B = T(X_2)$ for objects in the image of a compact tree, and $m = T(m_2)$ for the relevant morphism.

Essentially, the corollary allows the morphism b coming from injectivity $\prod_{\mathcal{U}} A_i \models T[c]$ of the tree with respect to the ultraproduct under the assignment c to be transferred to a large family of b_i morphisms that witness injectivity $A_i \models T[c_i]$ of the tree with respect to the large family of A_i objects and their corresponding c_i assignments.

Corollary 7.2.3.2. Let C and B be compact objects, and let $m : C \to B$.

Let $c : C \to \prod_{\mathcal{U}} A_i$ and assume there is $b : B \to \prod_{\mathcal{U}} A_i$ such that $bm = c$.

Then, there is $Z \in \mathcal{U}$ and families of morphisms

$$\langle c_i : C \to A_i\rangle_{i\in Z}$$

$$\langle b_i : B \to A_i\rangle_{i\in Z}$$

such that

$$\perp\!\!\!\perp_Z \langle c_i\rangle_{i\in Z} = c,$$
$$\perp\!\!\!\perp_Z \langle b_i\rangle_{i\in Z} = b,$$

and

$$\langle c_i\rangle_{i\in Z} = \langle b_i\, m\rangle_{i\in Z}.$$

Proof. By the first property of compactness, there are $X, Y \in \mathcal{U}$ such that

$$\underline{\amalg}_X \ \langle c_i \rangle_{i \in X} = c$$

and

$$\underline{\amalg}_Y \ \langle b_i \rangle_{i \in X} = b.$$

Since $bm = c$ then

$$\underline{\amalg}_Y \ \langle b_i \ m \rangle_{i \in X} \quad = \quad \underline{\amalg}_Y \ \langle b_i \rangle_{i \in X} \ m \quad = \quad \underline{\amalg}_X \ \langle c_i \rangle_{i \in X}$$

Hence, by the above lemma, there is some Z such that

$$\langle c_i \rangle_{i \in Z} \quad = \quad \langle b_i \ m \rangle_{i \in Z}.$$

Furthermore,

$$\begin{aligned} \underline{\amalg}_Z \ \langle c_i \rangle_{i \in Z} \quad &= \quad \underline{\amalg}_Z \ \pi_Z^X \ \langle c_i \rangle_{i \in X} \quad = \quad \underline{\amalg}_X \ \langle c_i \rangle_{i \in X} \quad = \quad c, \\ \underline{\amalg}_Z \ \langle b_i \rangle_{i \in Z} \quad &= \quad \underline{\amalg}_Z \ \pi_Z^Y \ \langle b_i \rangle_{i \in Y} \quad = \quad \underline{\amalg}_Y \ \langle b_i \rangle_{i \in Y} \quad = \quad b. \end{aligned}$$

□

The following simple lemma allows factorizations of morphisms into the ultraproduct to be done via "small enough" families

Lemma 7.2.3.3. Let $c : C \to \prod_{\mathcal{U}} A_i$ be a morphism from a (not necessarily compact) object $C \in \mathbb{C}$ into an ultraproduct.

Let $Z \in \mathcal{U}$.

Then, if for some $X' \in \mathcal{U}$ there is a factorization of c as $\underline{\amalg}_{X'} \ \langle c_i \rangle_{i \in X'}$, then there is a factorization $c = \underline{\amalg}_X \ \langle c_i \rangle_{i \in X}$, for some $X \in \mathcal{U}$ with $X \subseteq Z$.

Proof. Define $X = X' \cap Z$. Then in particular, $X \subseteq X'$ (and $X \in \mathcal{U}$), so,

$$\underline{\amalg}_{X'} \ \langle c_i \rangle_{i \in X'} \quad = \quad \underline{\amalg}_X \ \pi_X^{X'} \ \langle c_i \rangle_{i \in X'} \quad = \quad \underline{\amalg}_X \ \langle c_i \rangle_{i \in X}.$$

□

7.2.4 Examples of Compact Objects

- In any "algebraic" category, such as groups or rings, the compact objects are exactly the finitely presentable ones.
- The compact objects in the category of topological spaces are **not** the compact topological spaces, but have a more complicated description.

7.3 Proof of Łoś's Theorem

7.3.1 Statement of Theorems

These theorems are stated here almost exactly as they are in [AN78].

Theorem 7.3.1.1. Let $\mathbb{C}$ be a category with (small) products and (small) directed colimits. Let $T : \mathbb{T} \to \mathbb{C}$ be a compact tree. Let $\langle A_i \rangle_{i \in I}$ be a family of objects in $\mathbb{C}$ indexed by a set I and let $\mathcal{U} \subset \mathcal{P}(I)$ an ultrafilter on $\mathcal{P}(I)$.

Then the following two statements are equivalent.

1. $$\prod_{\mathcal{U}} A_i \models T$$

2. For some $W \in \mathcal{U}$,
$$A_i \models T \text{ for every } i \in W.$$

$\lhd$

The proof of this statement follows from Theorem 7.3.1.2. This theorem is first stated and used to prove the above theorem. Its proof then follows.

Theorem 7.3.1.2. Let $\mathbb{C}$ be a category with (small) products and (small) directed colimits, and let $T : \mathbb{T} \to \mathbb{C}$ be a compact tree. Let $\langle A_i \rangle_{i \in I}$ be a family of objects in $\mathbb{C}$ indexed by a set I and let $\mathcal{U} \subset \mathcal{P}(I)$ an ultrafilter on $\mathcal{P}(I)$.

Now, let $\langle k_i : T(\perp) \to A_i \rangle_{i \in J}$ be a cone of morphisms from $T(\perp)$ indexed by some set $J \subseteq I$.

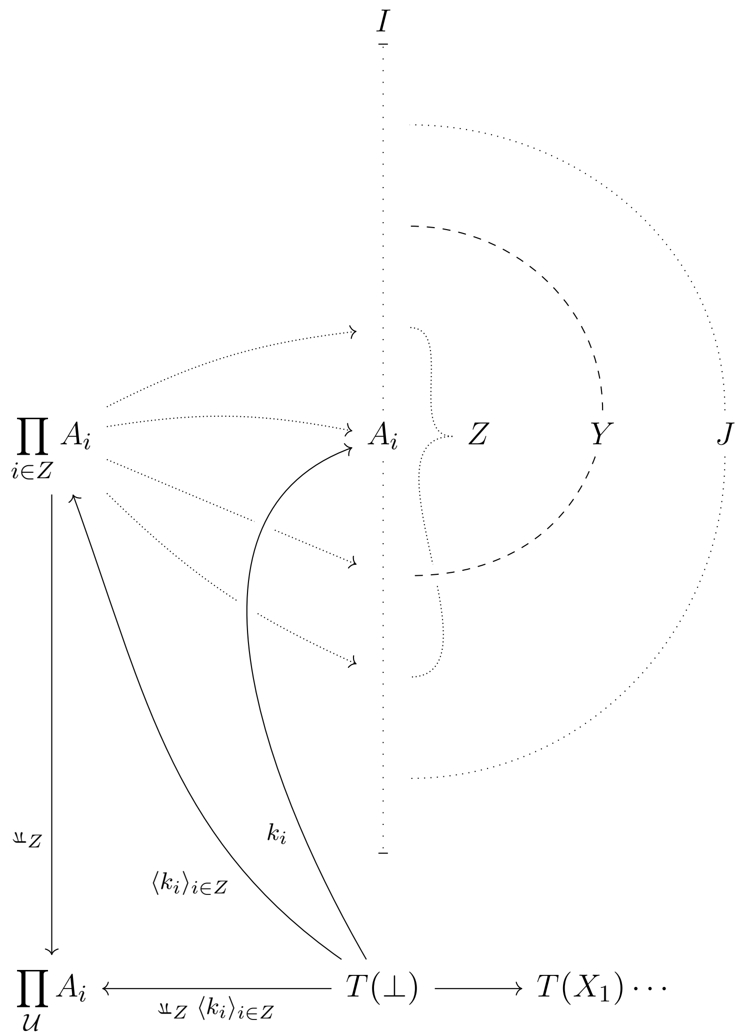
Then the following two statements are equivalent:

1. There is some $Z \in \mathcal{U}$ with $Z \subseteq J$ such that
$$\prod_{\mathcal{U}} A_i \models T[\, \Perp_Z \, \langle k_i \rangle_{i \in Z} \,]$$

2. There is some $Y \in \mathcal{U}$ with $Y \subseteq J$ such that
$$A_i \models T[k_i] \text{ for every } i \in Y.$$

Here, $\Perp_Z$ is the coprojection morphism from $\prod_{i \in Z} A_i$ into the ultraproduct and $\langle k_i \rangle_{i \in Z}$ is the morphism from $T(\perp)$ into the product $\prod_{i \in Z} A_i$ induced by the cone $\langle k_i \rangle_{i \in J}$.

◁



7.3.2 Intuition for Proof of Main Theorem 7.3.1.1 from Subsidiary Theorem 7.3.1.2

In a sense, the subsidiary theorem gives the necessary conditions for each "assignment of variables" from the image of the root $T(\perp)$ to the ultraproduct, and to each A_i in a large set. The main theorem is then just the statement for all such assignments at once.

For the forward direction, the necessary large set W of indices $i \in I$ for statement 2, for which A_i is injective, is given by the set of **all** $i \in I$ for which A_i is injective.

It must be shown this W is in $\mathcal{U}$. However, it is **not** in $\mathcal{U}$ if, and only if, its complement $J = I - W$ **is** in $\mathcal{U}$. Now, J consists of those i such that each A_i has a corresponding morphism $k_i : T(\perp) \to A_i$ for which A_i is **not** injective under the assignment k_i. Using the subsidiary theorem, this results in an index $i \in J$ (in particular, $i \in Y \subseteq J$) for which A_i is injective under the assignment k_i. This is a contradiction, and hence, W must be the large set.

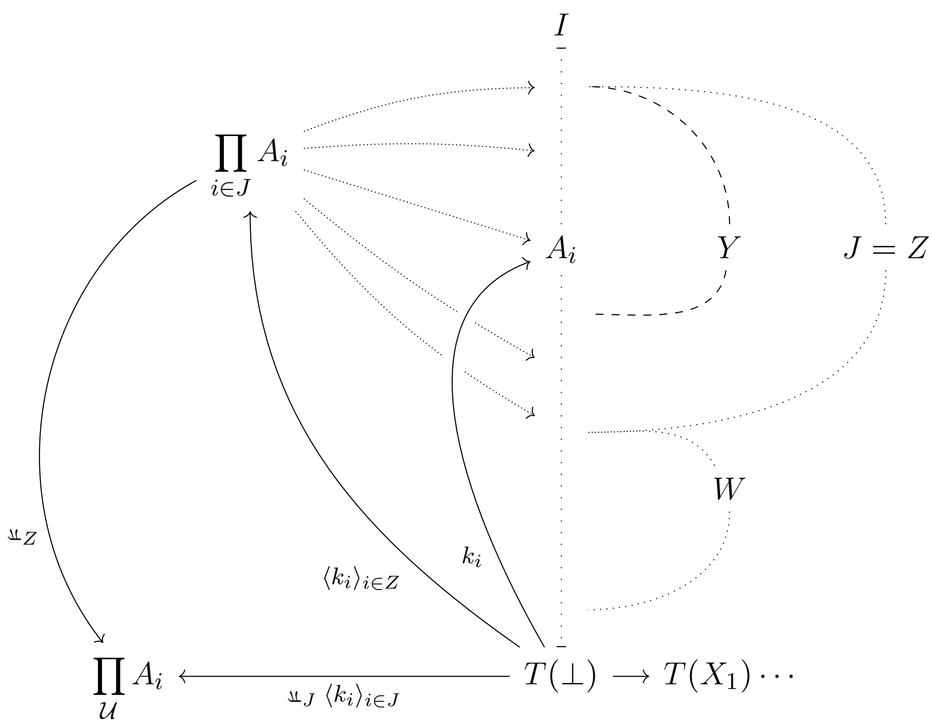
For the other direction, compactness of the tree allows us to factor each $k : T(\perp) \to \prod_{\mathcal{U}} A_i$ from the image of the root into the ultraproduct, through some product $\prod_{Y'} A_i$.

Now, we can intersect this Y' with W (i.e., with those i such that A_i is injective) from part 2 of the main theorem. This gives the Y we need to satisfy part 2 of the subsidiary theorem.

This implies part 1 of the subsidiary theorem, in such a way that the "variable assigment" morphism (in the square brackets) is actually k. Then, since k is arbitrary, this is exactly the meaning of injectivity of the ultraproduct.

7.3.3 Proof of Main Theorem 7.3.1.1 from Subsidiary Theorem 7.3.1.2

This proof is taken from [AN78].



Proof.
($\Rightarrow$) Suppose $\prod_{\mathcal{U}} A_i \models T$.

Define $J = \{i \in I : A_i \not\models T\}$.

Then, by definition of injectivity, there is a cone $\langle k_i : T(\perp) \to A_i \rangle_{i \in J}$ such that for every $i \in J$, it holds that $A_i \not\models T[k_i]$.

Also by injectivity, $\prod_{\mathcal{U}} A_i \models T[\ \perp\!\!\!\perp_J \langle k_i \rangle_{i \in J}\ ]$, assuming $J \in \mathcal{U}$.

Now, let $Z = J$. If $Z = J \in \mathcal{U}$, then since $J \subseteq J$, statement 1 of Theorem 7.3.1.2 is satisfied.

Then statement 2 is also satisfied, but this means there is some $i \in Y \subseteq J$, and thus in particular some $i \in J$ such that $A_i \models T[k_i]$ which contradicts the definition of J.

Hence $J \notin \mathcal{U}$, and so by the properties of ultrafilters, $I - J \in \mathcal{U}$.

And so, letting $W = I - J$ satisfies statement 2 of Theorem 7.3.1.1, namely $A_i \models T$ for every $i \in I - J$.

($\Leftarrow$) Suppose there is $W \in \mathcal{U}$ such that for every $i \in W$, it holds that $A_i \models T$.

Let $k : T(\perp) \to \prod_{\mathcal{U}} A_i$.

Then, since $T(\perp)$ is compact there is a set $Y' \in \mathcal{U}$ and morphism $k' : T(\perp) \to \prod_{i \in Y'} A_i$ such that $\perp\!\!\!\perp_{Y'} k' = k$. But then k' is the induced morphism $\langle k'_i m \rangle_{i \in Y'}$ for the **cone** $\langle k'_i : T(\perp) \to A_i \rangle_{i \in Y'}$.

$$
\begin{tikzcd}
 & \prod_{\mathcal{U}} A_i \\
T(\perp) \arrow[ur, "k"] \arrow[dr, "\langle k'_i \rangle_{i \in Y'}"'] & \\
 & \prod_{i \in Y'} A_i \arrow[uu, "\perp\!\!\!\perp_{Y'}"']
\end{tikzcd}
\qquad
\begin{tikzcd}
 & & \prod_{\mathcal{U}} A_i & & \\
T(\perp) \arrow[urr, "k"] \arrow[d, "\langle k'_i \rangle_{i \in Y'}"'] \arrow[drrrr, "\langle k_i \rangle_{i \in Y}"] & & & & \\
\prod_{i \in Y'} A_i \arrow[uurr, "\perp\!\!\!\perp_{Y'}"] \arrow[rrrr, "\pi^{Y'}_Y"'] & & & & \prod_{i \in Y} A_i \arrow[uull, "\perp\!\!\!\perp_{Y}"']
\end{tikzcd}
$$

Now, for any $Y \subseteq Y'$, there is projection

$$\pi^{Y'}_Y : \prod_{i \in Y'} A_i \to \prod_{i \in Y} A_i$$

and if $Y \in \mathcal{U}$, then by commutativity of the coprojection cone into the ultra-product, $\perp\!\!\!\perp_Y \pi^{Y'}_Y = \perp\!\!\!\perp_{Y'}$.

Now define $Y = Y' \cap W$. Then by the properties of ultrafilters, $Y \in \mathcal{U}$. (Also $Y \subseteq Y'$ so the projection exists). Then define

$$\langle k_i \rangle_{i \in Y} = \pi^{Y'}_Y \langle k'_i \rangle_{i \in Y'}$$

which is the same as defining each $k_i = k'_i$. By the above, this results in $\Perp_Y \langle k_i \rangle_{i \in Y} = k$.

By assumption, and the definition of injectivity, for every k_i, it is true that $A_i \models T[k_i]$. Hence part 2 of Lemma 7.3.1.2 is satisfied. So part 1 holds also.

Thus there is a $Z \in \mathcal{U}$ with $Z \subseteq J$ such that

$$\prod_{\mathcal{U}} A_i \models T[\ \Perp_Z \langle k_i \rangle_{i \in Z}\ ].$$

But again by the commutativity of the ultraproduct cone, $\Perp_Z \langle k_i \rangle_{i \in Z} = k$, and so, since k was arbitrary,

$$\prod_{\mathcal{U}} A_i \models T.$$

$\square$

7.3.4 Intuition for Subsidiary Theorem 7.3.1.2

This subsidiary theorem actually bears the brunt of the proof.

The basic proof is taken from [AN78], although there are some changes to the structuring and diagrams of the proof, and more detail has been provided.

The proof uses well-founded induction (§1.4) on depth-two subtrees. It is assumed that the theorem holds for all depth-two subtrees, and it is shown that this implies it holds for the entire tree. The motivation for using **depth two** subtrees (instead of simply subtrees) is because of the defintion of injectivity containing alternating forall-exists.

Remark. The induction does not need a base case, or, rather, the base case is implicit in the statement. If a given tree has **no** depth-two subtrees, then the statement vacuously holds for all its depth-two subtrees, and the proof is still valid for this case. A tree with no depth-two subtrees hence satisfies the given conditions.

For example, if a tree has no sub-trees at all, then all its subtrees vacuously satisfy the condition of the theorem, so by the implication condition in the induction, the (singleton) tree satisfies the condition of the theorem.

Similarly, even if a tree has only depth (at most) **one** subtrees, then all of those depth-one subtrees must be (singleton) trees without subtrees of their own (since if they had any subtrees, their depth-one subtrees would be depth-two subtrees of the super-tree). Hence they satisfy the condition of the theorem. So by implication (induction), the tree satisfies the condition. $\triangleleft$

For the forward direction, the (forward direction of the) theorem is assumed for all depth-two subtrees (trees whose root is a layer **three** object of the main tree).

Given that the ultraproduct is injective under the "assignment" of the specific morphism $\pitchfork_Z \langle k_i \rangle_{i \in Z}$ factoring through $\prod_Z A_i$, some large-indexed family of the A_i must be found for which each particular assignment k_i – the components of the morphism into the product – is also an assignment making A_i injective.

Injectivity provides a layer 2 object $T(X_2)$ (and corresponding morphisms) such that every layer 3 object $T(X_3)$ is injective.

This morphism from the layer 2 object into the ultraproduct coming from injectivity can be factored through one of the large products (say indexed by W) using compactness, say $\pitchfork_W \langle k'_i \rangle_{i \in W} : T(X_2) \to \prod_{\mathcal{U}} A_i$. In particular, also, each projection $k_i : T(\perp) \to A_i$ for $i \in W$ can be factored via $T(X_2)$ as $k'_i T(m_2)$.

The subset $Y \subseteq W$ of A_is injective under the k_i assignments provides the Y necessary for the proof. It must just be shown to be in the ultrafilter $\mathcal{U}$.

If it weren't in the ultrafilter, this would lead to a contradiction: Since this is the set of A_i which are **not** injective under the assignments k_i, each has a layer 3 object in the tree, with corresponding morphism into that A_i which fails injectivity. Because there are finitely many layer 3 objects, the partition condition implies there is a single layer 3 object X_3^r which fails for a large subfamily of the A_i, indexed, say, by Q.

Using the contrapositive of the induction step on X_3^r, failure of part 2 implies failure of part 1, which leads to the non-existence of a working Y (for part 1 of the theorem). But the Y from above **does** work. Hence, this is a contradiction, and so Y is in the ultrafilter.

For the reverse direction, it is again possible to use the partition condition to find a single working object X_2 (this time of layer 2) such that every suitable subtree $T_{X_3^j}$ and corresponding morphism q_j satisfies the injectivity conditions.

Now, based on the above, a specific pair of morphisms $\bar{k}$ and $\bar{k}'$ from $T(\perp)$ and $T(X_2^r)$ respectively into the ultrafilter are defined based on this.

These are factored out to certain (large) products. The inductive assumption on subtrees is used to get the relevant morphisms into the ultraproduct, and the lemmas for compact objects provide the existence of a (small enough) product necessary for everything to commute.

7.3.5 Proof of Subsidiary Theorem 7.3.1.2

Proof of Theorem 7.3.1.2.
Suppose the assumptions of the theorem hold: There is family of objects in $\mathbb{C}$ given by $\langle A_i\rangle_{i\in I}$ and indexed by set I. There is ultrafilter $\mathcal{U} \subset \mathcal{P}(I)$ on $\mathcal{P}(I)$. There is cone $\langle k_i : T(\perp) \to A_i\rangle_{i\in J}$ of morphisms from $T(\perp)$ indexed by set $J \subseteq I$. The particular J here is not important, so from this point on we use it as a variable in the induction argument.

($\Rightarrow$) Assume the implication (1) $\Rightarrow$ (2) holds for every subtree of T with a root coming from a layer three object X_3 of $\mathbb{T}$,

$$T_{X_3}(X_3) : \mathbb{T}_{X_3} \to \mathbb{C}, \quad \text{for } X_3 \in \mathbb{T}, \quad \mathrm{Layer}(X_3) = 3,$$

and for any cone $\langle T(X_3) \to A_i\rangle_{i\in J}$, where $J \subseteq I$.

(Bear in mind that X_3 is the root of the subtree $X_3 = \perp_{T_{X_3}}$)

Now assume (1) holds for T, so there is some $Z \in \mathcal{U}$ with $Z \subseteq J$ such that

$$\prod_{\mathcal{U}} A_i \models T[\, \perp\!\!\!\perp_Z \langle k_i\rangle_{i\in Z}\,]. \tag{7.3.1}$$

Define

$$\bar{k} = \perp\!\!\!\perp_Z \langle k_i\rangle_{i\in Z}. \tag{7.3.2}$$

Then, by definition of injectivity, there is a layer 2 object $X_2 \in \mathbb{T}$ and morphisms $m_2 : \perp \to X_2$ and $\bar{k}' : T(X_2) \to \prod_{\mathcal{U}} A_i$ such that

$$\bar{k} = \bar{k}'\, T(m_2), \tag{7.3.3}$$

and furthermore such that for any layer 3 object $X_3 \in \mathbb{T}$ and morphisms $m_3 : X_2 \to X_3$ and $q : T(X_3) \to \prod_{\mathcal{U}} A_i$ such that

$$\bar{k}' = q\, T(m_3). \tag{7.3.4}$$

it holds that

$$\prod_{\mathcal{U}} A_i \models T_{X_3}[q]. \tag{7.3.5}$$

$$\begin{tikzcd}
T(X_3) \arrow[dr, "q"] & \\
T(X_2) \arrow[u, "T(m_3)"] \arrow[r, "\bar{k}'"] & \prod_{\mathcal{U}} A_i \\
T(\perp) \arrow[u, "T(m_2)"] \arrow[ur, "\bar{k}"] \arrow[r, "\langle k_i\rangle_{i\in Z}"] \arrow[dr, "k_i"'] & \prod_{i\in Z} A_i \arrow[u, "\perp\!\!\!\perp_Z"'] \arrow[d, "\pi_i"] \\
& A_i
\end{tikzcd}$$

Now, since T is compact, then in particular, $T(\perp)$ and $T(X_2)$ are compact.

By Corollary 7.2.3.2 and Lemma 7.2.3.3 there is a set $W \in \mathcal{U}$ with $W \subseteq Z$ and family $\langle k_i' \rangle_{i \in W}$ such that

$$\perp\!\!\!\perp_W \ \langle k_i \rangle_{i \in W} = \bar{k},$$

$$\perp\!\!\!\perp_W \ \langle k_i' \rangle_{i \in W} = \bar{k}',$$

and

$$\langle k_i \rangle_{i \in W} = \langle k_i' \, T(m_2) \rangle_{i \in W}. \tag{7.3.6}$$

Remark. Since it is already true that

$$\perp\!\!\!\perp_Z \ \langle k_i \rangle_{i \in Z} = \bar{k}$$

then this can be taken as the factorization of $\bar{k}$ guaranteed by compactness. It then follows that

$$\perp\!\!\!\perp_W \ \langle k_i \rangle_{i \in W} \quad = \quad \perp\!\!\!\perp_Z \pi_W^Z \ \langle k_i \rangle_{i \in W} \quad = \quad \perp\!\!\!\perp_Z \ \langle k_i \rangle_{i \in Z} \quad = \quad \bar{k}.$$

$\lhd$

Now, define

$$Y = \{ i \in W : A_i \models T[k_i] \}. \tag{7.3.7}$$

It is shown that $Y \in \mathcal{U}$ and hence that this is the necessary set to complete the proof.

Suppose, for contradiction, that $Y \notin \mathcal{U}$. Then, $I - Y \in \mathcal{U}$ and so $W \cap (I - Y) = W - Y \in \mathcal{U}$. This set may be written

$$W - Y = \{ i \in W : A_i \not\models T[k_i] \}.$$

And, by Equation 7.3.6, $k_i = k_i' \, T(m_2)$ for $i \in W$.

Now, by definition of $A_i \not\models T[k_i]$, there is, for each $i \in W - Y$, a layer 3 object $X_3^{r(i)}$ and morphisms $m_3^{r(i)} : X_2 \to X_3^{r(i)}$ and $q_i : T(X_3^{r(i)}) \to A_i$ such that

$$q_i \, T(m_3^{r(i)}) = k_i' \tag{7.3.8}$$

and

$$A_i \not\models T_{X_3^{r(i)}}[q_i].$$

$$\begin{tikzcd}
T(X_2) \arrow[d, "T\left(m_3^{r(i)}\right)"'] \arrow[dr, "k_i'"] & \\
T(X_3) \arrow[r, "q_i"] & A_i
\end{tikzcd}$$

Let n denote the (finite) number of layer 3 objects X_3^i.

Now, $r : W - Y \to \mathbb{N}$ is a function taking on values less than or equal to n. (So, in fact, $r : W - Y \to (n+1)$.) Then by the partition condition of §3.4 applied to $\mathcal{U}$, there is subset $Q \subseteq W - Y$ with $Q \in \mathcal{U}$, and number, also called r, such that $r = r(i)$ for every $i \in Q$.

Since for all $i \in Q$,

$$A_i \not\models T_{X_3^r}[q_i]$$

then for each $Q' \subseteq Q$, with $Q' \in \mathcal{U}$, there is $i \in Q'$ such that

$$A_i \not\models T_{X_3^r}[q_i].$$

Now, by the contrapositive of the assumption for the induction, letting $J = Q$ implies for every $Q' \subseteq Q$, with $Q' \in \mathcal{U}$, that

$$\prod_{\mathcal{U}} A_i \not\models T_{X_3^r}[\Perp_{Q'} \langle q_i \rangle_{i \in Q'}]. \tag{7.3.9}$$

and, in particular, such a Q' exists, e.g., by letting $Q' = Q$, so this is not just a trivially true universal statement.

But, by Equation 7.3.8, for any such Q' it holds that

$$\begin{aligned} \bar{k}' &= \Perp_W \langle k_i' \rangle_{i \in W} \\ &= \Perp_{Q'} \pi_{Q'}^W \langle k_i' \rangle_{i \in W} \\ &= \Perp_{Q'} \langle k_i' \rangle_{i \in Q'} \\ &= \Perp_{Q'} \langle q_i \rangle_{i \in Q'} T(m_3^r) \end{aligned}$$

and so, letting $q = \Perp_{Q'} \langle q_i \rangle_{i \in Q'}$ in Equation 7.3.4 and Statement 7.3.5 yields

$$\prod_{\mathcal{U}} A_i \models T_{X_3^r}[\Perp_{Q'} \langle q_i \rangle_{i \in Q'}]$$

which contradicts Statement 7.3.9.

$$\begin{tikzcd}
 & & \prod_{\mathcal{U}} A_i & \\
T(X_2) \arrow[urr, "\bar{k}'"] \arrow[d, "T(m_3)"'] \arrow[drr, "\langle k_i' \rangle_{i \in W}"] \arrow[drrr, "\langle k_i' \rangle_{i \in Q'}"] & & & \\
T(X_3) \arrow[uurr, "q"] \arrow[rrr, bend right, "\langle q_i \rangle_{i \in Q'}"'] & & \prod_{i \in W} A_i \arrow[uu, "\Perp_W"'] \arrow[r, "\pi_{Q'}^W"'] & \prod_{i \in Q'} A_i \arrow[uul, "\Perp_{Q'}"']
\end{tikzcd}$$

Thus $Y \in \mathcal{U}$.

($\Leftarrow$) Suppose the implication (2) $\Rightarrow$ (1) holds for every subtree of T with a root coming from a layer three object X_3 of $\mathbb{T}$,

$$T_{X_3}(X_3) : \mathbb{T}_{X_3} \to \mathbb{C}, \qquad \text{for } X_3 \in \mathbb{T}, \quad \text{Layer}(X_3) = 3,$$

and for any cone $\langle T(X_3) \to A_i \rangle_{i \in J}$, where $J \subseteq I$.

It is then shown that (2) $\Rightarrow$ (1) for T also.

Assume (2) holds for T, so there is $Y \in \mathcal{U}$ with $Y \subseteq J$ such that

$$A_i \models T[k_i] \text{ for every } i \in Y.$$

As in the diagram,

$$\begin{tikzcd}
 & A_i & & \\
T(\perp_T) \arrow[ur, "k_i"] \arrow[r, "T(m_{r(i)})"] & T(X_2^{r(i)}) \arrow[u, "k_i'"] \arrow[r, "T(m_j^{r(i)})"] & T(X_3^{r(i),j}) \arrow[ul, "q_j^i"'] & \cdots
\end{tikzcd}$$

by the definition of injectivity, for each $i \in Y$ there exists a layer 2 object $X_2^{r(i)}$ and morphisms $m_{r(i)} : \perp_T \to X_2^{r(i)}$ and $k_i' : T(X_2^{r(i)}) \to A_i$ such that

$$k_i = k_i' \, T(m_{r(i)}) \tag{7.3.10}$$

and, given such a morphism $m_{r(i)}$, then for every layer 3 object $X_3^{r(i),j}$ and morphisms $m_j^{r(i)} : X_2^{r(i)} \to X_3^{r(i),j}$ and $q_j^i : T(X_3^{r(i),j}) \to A_i$ for which $q_j^i T(m_j^{r(i)}) = k_i'$, it holds that

$$A_i \models T_{X_3^{r(i),j}}[q_i]. \tag{7.3.11}$$

Let n denote the (finite) number of layer 2 objects X_2^i.

Now, $r : Y \to \mathbb{N}$ is a function taking on values less than or equal to n. (So, in fact, $r : Y \to (n+1)$.) Then by the partition condition of §3.4 applied to $\mathcal{U}$, there is subset $Z' \in \mathcal{U}$ and number, also called r, such that $r = r(i)$ for every $i \in Z'$.

In particular, since $Z = Z' \cap Y$ is in $\mathcal{U}$, then for this $Z \subseteq Y$, also

$$r = r(i), \qquad \text{for every } i \in Z. \tag{7.3.12}$$

This allows us to swap some of the quantification in the above characterization of injectivity. In particular, now there exists a single X_2 and morphism $m_r : \perp \to X_2$ such that for every $i \in Z$, it holds that $k_i' \, T(m_r) = k_i$ and, for every X_3^j and $m_j : X_2 \to X_3^j$ and $q_j : T(X_3^j) \to A_i$ with $q_j \, T(m_j) = k_i'$, it holds that

$$A_i \models T_{X_3^j}[q_j]. \tag{7.3.13}$$

Now define morphisms

$$\bar{k} = \text{ш}_Z \, \langle k_i \rangle_{i\in Z} \quad : \quad T(\perp) \to \prod_{\mathcal{U}} A_i \tag{7.3.14}$$

$$\bar{k}' = \text{ш}_Z \, \langle k_i' \rangle_{i\in Z} \quad : \quad T(X_2^r = X_2) \to \prod_{\mathcal{U}} A_i \tag{7.3.15}$$

It is shown that the morphism $\bar{k}'$ is the morphism that exists satisfying the requirements for injectivity of $\bar{k}$ in

$$\prod_{\mathcal{U}} A_i \models T\,[\bar{k}]\,.$$

Since, by Equations 7.3.10 and 7.3.12, for every $i \in Z$ it holds that $k_i = k_i' \, T(m_r)$, then by the properties of products,

$$\langle k_i \rangle_{i\in Z} = \left\langle k_i' \right\rangle_{i\in Z} \, T(m_r)$$

and so,

$$\text{ш}_Z \, \langle k_i \rangle_{i\in Z} = \text{ш}_Z \, \left\langle k_i' \right\rangle_{i\in Z} \, T(m_r),$$

and hence, by Equations 7.3.14 and 7.3.15, $\bar{k}'$ satisfies the required commutativity

$$\bar{k} = \bar{k}' \, T(m_r). \tag{7.3.16}$$

$$\begin{tikzcd}
 & \prod_{\mathcal{U}} A_i & \\
T(\perp) \arrow[ur, "\bar{k}"] \arrow[d, "T(m_r)"'] \arrow[rr, "\langle k_i \rangle_{i\in Z}"] & & \prod_{i\in Z} A_i \arrow[ul, "\text{ш}_Z"'] \\
T(X_2^r) \arrow[uur, "\bar{k}'"] \arrow[urr, "\langle k_i' \rangle_{i\in Z}"'] & &
\end{tikzcd}$$

Let X_3 be a layer 3 object in the tree with morphism $m : X_2^r \to X_3$, and let $q : T(X_3) \to \prod_{\mathcal{U}} A_i$ be a morphism commuting by

$$\bar{k}' = q\, T(m). \tag{7.3.17}$$

Then it is shown that

$$\prod_{\mathcal{U}} A_i \models T_{X_3}[q].$$

The diagram below denotes some of the equations and definitions which follow.

By Corollary 7.2.3.2 and Lemma 7.2.3.3 there is a set $W \in \mathcal{U}$ with $W \subseteq Z$ and family $\langle q_i \rangle_{i\in W}$ such that

$$\veebar_W \langle k_i' \rangle_{i\in W} = \bar{k}', \tag{7.3.18}$$

$$\veebar_W \langle q_i \rangle_{i\in W} = q, \tag{7.3.19}$$

and

$$\langle k_i' \rangle_{i\in W} = \langle q_i\, T(m) \rangle_{i\in W}. \tag{7.3.20}$$

$$\begin{tikzcd}
 & \prod_{\mathcal{U}} A_i & \\
T(X_2^r) \arrow[ur, "\bar{k}'"] \arrow[d, "T(m)"'] \arrow[drr, "\langle k_i' \rangle_{i\in W}"] & & \\
T(X_3) \arrow[uur, "q"] \arrow[rr, "\langle q_i \rangle_{i\in W}"] & & \prod_{i\in W} A_i \arrow[uul, "\veebar_W"']
\end{tikzcd}$$

Now, for each $i \in W$, it holds that $k_i' = q_i\, T(m)$ and so, by Statement 7.3.11 and Equation 7.3.12,

$$A_i \models T_{X_3}[q_i].$$

Hence, statement (2) of the theorem holds for the subtree with both $J = W$ and $Y = W$, i.e., there is tree T_{X_3} and cone $\langle q_i : T_{X_3}(X_3) \to A_i \rangle_{i\in W}$, and $W \in \mathcal{U}$ (where $W \subseteq W$) such that

$$A_i \models T_{X_3}[q_i] \text{ for every } i \in W.$$

So, by the assumption that the theorem holds for the subtrees, then there is $Q \subseteq W$ with $Q \in \mathcal{U}$ such that

$$\prod_{\mathcal{U}} A_i \models T_{X_3}[\veebar_Q \langle q_i \rangle_{i\in Q}],$$

and hence, by Equation 7.3.19

$$\prod_{\mathcal{U}} A_i \models T_{X_3}[q].$$

Since this holds for arbitrary subtree T_{X_3} at any X_3 with morphism $q : T_{X_3}(X_3) \to \prod_{\mathcal{U}} A_i$ provided that q satisfies the commutativity property $q\,T(m) = \bar{k}'$, then this is exactly the last property required to show that

$$\prod_{\mathcal{U}} A_i \models T[k_i].$$

□

Corollary 7.3.5.1. Łoś's Theorem applied to the category of commutative rings implies that every ultraproduct of fields is a field.

Proof. This is proven is Section 7.4. □

7.4 Fields in the Category of Commutative Rings

The intention of this section is, given a commutative ring K with unity, to classify whether K is a field based on the existence of specific morphisms into K.

The author is not aware of any source which contains this result, and especially none which relate this fact to the general category-theoretic version Łoś's Theorem.

The algebra used in this section is quite basic and may be found in many standard textbooks on algebra or commutative algebra. See, for example, [AM69].

Let $\mathbb{C}$ be the category of commutative rings with unity, and such that $0 \neq 1$.

The ring $\mathbb{Z}[X]$ is the ring of polynomials in X with coefficients in the integers. This has the property that for any ring R, for each element $r \in R$ there is a unique homomorphism $f_r : \mathbb{Z}[X] \to R : X \mapsto r$.

For any commutative ring R and any subset $Q \subseteq R$ not containing nilpotent elements, there exists the ring $R[Q^{-1}]$ and homomorphism $f : R \to R[Q^{-1}]$, called the localization of R at Q. This is the universal ring such that every element of Q is invertible, i.e., such that for any $q \in Q$, the image $f(q)$ is a unit in $R[Q^{-1}]$. It is universal in the sense that for any other ring S and homomorphism $g : R \to S$ with $g(q)$ a unit for every $q \in Q$, there is a unique $h : R[Q^{-1}] \to S$ with $g = hf$.

The localization of $\mathbb{Z}[X]$ at $\{X\}$ is a ring of polynomials with coefficients in the integers, but where the powers of X are allowed to be negative. I.e., of the form: $a_{-m}X^{-m} + a_{-m+1}X^{-m+1} + \cdots + a_{n-1}X^{n-1} + a_nX^n$, where $a_i \in \mathbb{Z}$ for all i. Denote this $\mathbb{Z}[X][X^{-1}]$.

Consider the following tree of rings with corresponding maps

$$
\begin{tikzcd}
 & \mathbb{Z}[X][X^{-1}] \\
\mathbb{Z}[X] \arrow[ur, "f"] \arrow[dr, "g"'] & \\
 & \mathbb{Z}
\end{tikzcd}
$$

$$
\begin{aligned}
f :&\quad \mathbb{Z}[X] \to \mathbb{Z}[X][X^-1] &&: X \mapsto X \\
g :&\quad \mathbb{Z}[X] \to \mathbb{Z} &&: X \mapsto 0.
\end{aligned}
$$

Theorem 7.4.0.1. For any commutative ring K, K is a field if and only if, for every morphism $h_1 : \mathbb{Z}[X] \to K$, there exists either a commuting morphism $h_2 : \mathbb{Z} \to K$ or commuting morphism $h_3 : \mathbb{Z}[X][X^{-1}] \to K$ ($h_1 = h_2g$ or $h_1 = h_3f$).

$$
\begin{tikzcd}
 & K & \\
 & & \mathbb{Z}[X][X^{-1}] \arrow[ul, "h_3"'] \\
\mathbb{Z}[X] \arrow[uur, "h_1"] \arrow[urr, "f"] \arrow[drr, "g"'] & & \\
 & & \mathbb{Z} \arrow[uuul, "h_2"]
\end{tikzcd}
$$

Proof. ($\Rightarrow$) Assume K is a field.

Any morphism from $\mathbb{Z}[X]$ is determined by where it maps X, so provided that there are two homomorphisms mapping X to the same place, they must be equal.

Let $h_1(X) = X_K$.

There is a unique map from $\mathbb{Z}$ to K. Now, $h_2g(X) = h_2(0) = 0$ so that if $X_K = 0$, then $h_1 = h_2g$.

Assume $X_K \neq 0$. The map h_3 is determined by where it maps X, since $1 = h_3(1) = h_3(XX^{-1}) = h_3(X)h_3(X^{-1})$, so that the inverse of $h_3(X)$ must exist and X^{-1} must be mapped to it for h_3 to be a homomorphism. However, provided that X_K has an inverse (which it does since it is non-zero and K is a field), then letting $h_3(X) = X_K$ means that h_3 can be completed to a homomorphism, which furthermore commutes with h_1.

($\Leftarrow$) Let $x \in K$, and define h_1 such that $h_1(X) = x$. If $x \neq 0$, then there cannot be a map h_2 with $h_1 = h_2 g$.

Hence there is a map h_3 with $h_1 = h_3 f$. This means that $X^{-1} \in \mathbb{Z}[X][X^{-1}]$ is mapped to $h_3(X^{-1})$ and that $xh_3(X^{-1}) = h_1(X)h_3(X^{-1}) = h_3(X)h_3(X^{-1}) = h_3(XX^{-1}) = h_3(1) = 1 \in K$. Hence x has inverse $h_3(X^{-1})$.

So being a field is equivalent to injectivity with respect to this tree. □

Remark. Since all of the objects in this tree are finitely presentable, then this tree is compact.

Corollary 7.4.0.2. A category-theoretic ultraproduct of fields in the category of rings is itself a field.

Proof. This follows directly from the generalized version of Łoś's Theorem applied to the category of rings, and the use of the above tree. □

Chapter 8

Ultrafilters and Ultraproducts Using Codensity Monads

8.1 Density

The concept of density is standard in category theory. Its definition may be found in [ML98, Chapter 10, §6]. Essentially, a functor $F : \mathbb{C} \to \mathbb{D}$ is **dense** in $\mathbb{D}$ if every object of $\mathbb{D}$ is a **colimit** of objects in the image of F.

The concept of codensity is, of course, dual to the concept of density. Hence, F is dense in $\mathbb{D}$ if every object of $\mathbb{D}$ is a **limit** of objects in the image of F.

Below, the codensity monad of the inclusion functor

$$I : \mathbf{FinSet} \to \mathbf{Set}$$

is defined. This monad measures how far I is from being codense, in the sense that if I were codense, the codensity monad would be trivial.

As it turns out, (the functor of) the codensity monad for this inclusion corresponds exactly to the ultrafilter functor on **Set**. If it were trivial, all ultrafilters would be principal.

So, the existence of non-principal ultrafilters is in some sense a direct consequence of the fact that not every set is a limit of finite sets.

8.2 Induced Hom Functor

Let $G : \mathbb{B} \to \mathbb{A}$ be a functor. Then there is an induced functor $\mathrm{Hom}(-, G)$.

Definition 8.2.0.1. The functor

$$\mathrm{Hom}(-, G) : \mathbb{A} \to \left(\mathbf{Set}^{\mathbb{B}}\right)^{\mathrm{op}}$$

is defined as follows (see diagrams at end of definition):

An object $A \in \mathbb{A}$ is sent to a (covariant) functor

$$\mathrm{Hom}(-, G)(A) = \mathrm{Hom}(A, G(-)) : \mathbb{B} \to \mathbf{Set}$$

which itself sends an object $B \in \mathbb{B}$ to the hom set

$$\mathrm{Hom}(A, G(-))(B) = \mathrm{Hom}_{\mathbb{A}}(A, G(B))$$

and sends a morphism $g : B \to B'$ to the left composition with $G(g)$ set-map between hom sets

$$\begin{array}{rrcr} \mathrm{Hom}(A, G(-))(g) : & \mathrm{Hom}_{\mathbb{A}}(A, G(B)) & \to & \mathrm{Hom}_{\mathbb{A}}(A, G(B')) \\ & h : A \to G(B) & \mapsto & G(g)h : A \to G(B'). \end{array}$$

For objects $A, A' \in \mathbb{A}$, the morphism $f : A \to A'$ is sent to a natural transformation in the opposite direction

$$\mathrm{Hom}(-, G)(f) = \mathrm{Hom}(f, G(-)) : \mathrm{Hom}(A', G(-)) \Rightarrow \mathrm{Hom}(A, G(-))$$

whose component at $B \in \mathbb{B}$ is given by the right composition set map

$$\begin{array}{rrcr} (\mathrm{Hom}(f, G(-))_B : & \mathrm{Hom}_{\mathbb{A}}(A', G(B)) & \to & \mathrm{Hom}_{\mathbb{A}}(A, G(B)) \\ & k : A' \to G(B) & \mapsto & kf : A \to G(B). \end{array}$$

$$\begin{tikzcd} & A \arrow[dl, "h"'] \arrow[dr, "G(g)h"] & \\ G(B) \arrow[rr, "G(g)"'] & & G(B') \end{tikzcd} \qquad \begin{tikzcd} A \arrow[rr, "f"] \arrow[dr, "kf"'] & & A' \arrow[dl, "k"] \\ & G(B) & \end{tikzcd}$$

◁

8.2.1 Codensity

Lemma 8.2.1.1. The functor G is **codense** if, and only if, $\mathrm{Hom}(-, G)$ is full and faithful. ◁

Proof. See Proposition 2 of [ML98, Chapter 10, §6]. □

Assume that $\mathbb{B}$ is essentially small (equivalent to a small category) and that $\mathbb{A}$ has all small limits.

Then $\mathrm{Hom}(-,G)$ has a right adjoint, which we shall denote $\mathrm{Nat}(-,G)$. So $\mathrm{Hom}(-,G) \dashv \mathrm{Nat}(-,G)$.

This is discussed in more detail in the following two sections.

Remark. The notation is motivated by the fact that in a specific case of interest, $\mathrm{Nat}(F,G)$ will be the set of natural transformations from F to G. ◁

8.3 Induced Nat Functor

The "Nat" functor of this section (§8.3) is one of the two "Hom" functors in [Lei13]. This section follows the definitions of that paper, but they are described in much more detail, and many of the proofs omitted in that paper are given here.

8.3.1 Nat Functor as Limit

The functor

$$\mathrm{Nat}(-,G) : \left(\mathbf{Set}^{\mathbb{B}}\right)^{\mathrm{op}} \to \mathbb{A}$$

is defined as follows:

For an object $F \in \mathbf{Set}^{\mathbb{B}}$, (i.e., a functor $F : \mathbb{B} \to \mathbf{Set}$),

$$\mathrm{Nat}(-,G)(F) = \mathrm{Nat}(F,G) = \varprojlim_{B\in\mathbb{B},\, x\in F(B)} GP_F(x,B)$$

In more detail:

Consider the comma category $(* \downarrow F)$ of 'elements of F' (where $*$ denotes the one-element set). Objects of this category are pairs (x,B) where $B \in \mathbb{B}$ and $x : * \to F(B)$. The morphism x can be seen as an element $x \in F(B)$. A morphism $(x,B) \to (x',B')$ is a morphism $(f : B \to B')$ such that $F(f) : F(B) \to F(B)'$ such that $x \mapsto x'$ in the sense that $xF(f) = x'$ or, put another way $F(f)(x) = x'$.

Then there is a projection functor

$$(* \downarrow F) \xrightarrow{P_F} \mathbb{B}$$

$$\begin{tikzcd} (x,B) \arrow[d, "g"] \arrow[r, phantom, "\longmapsto"] & B \arrow[d, "g"] \\ (x',B') & B' \end{tikzcd}$$

Remark. In particular, the image of g under the projection P_F has the property that its image $F(g)$ under F has to take x to x'. $\lhd$

There is a composition of the functors

$$(* \downarrow F) \xrightarrow{P_F} \mathbb{B} \xrightarrow{G} \mathbb{A}$$

Then the vertex of the limiting cone of this composition is

$$\mathrm{Nat}(F,G) = \varprojlim GP_F$$

For a natural transformation $\alpha : F' \Rightarrow F$, its image

$$\mathrm{Nat}(\alpha, G) : \mathrm{Nat}(F,G) \to \mathrm{Nat}(F',G)$$

i.e.,

$$\mathrm{Nat}(\alpha, G) : \varprojlim GP_F \to \varprojlim GP_{F'}$$

is defined as follows:

The limit $\varprojlim GP_{F'}$ has a limiting cone κ with components $\kappa_{x,B}$. This cone is universal with the property for any $g : B \to B'$ that $G(g)\kappa_{x,B} = \kappa_{F'(g)(x),B'}$.

Given $x \in F'(B)$, one obtains $\alpha_B(x) \in F(B)$. Now consider specifically the morphisms $\kappa_{\alpha_B(x),B}$ with codomains $G(B)$.

Then, since α is a natural transformation from F' to F, given $g : B \to B'$ for any x it holds that

$$\alpha_{B'}(F'(g)(x)) = F(g)(\alpha_B(x)) \tag{8.3.1}$$

The limit $\varprojlim GP_F$ has a limiting cone λ with components $\lambda_{y,B}$. This cone has the property for any $g : B \to B'$ that

$$G(g)\lambda_{y,B} = \lambda_{F(g)(y),B'} \tag{8.3.2}$$

then define

$$\gamma_{x,B} = \lambda_{\alpha_B(x),B}. \tag{8.3.3}$$

By equations 8.3.1, 8.3.2 and 8.3.3:

$$G(g)\gamma_{x,B} = G(g)\lambda_{\alpha_B(x),B} = \lambda_{F(g)(\alpha_B(x)),B'} = \lambda_{\alpha_{B'}(F'(g)(x)),B'} = \gamma_{F'(g)(x),B'}$$

And so, the $\gamma_{\alpha_B(x),B}$ form a cone over the diagram of the limit $\mathrm{Nat}(F', G)$. Hence there is a unique morphism from $\mathrm{Nat}(F, G)$ to $\mathrm{Nat}(F', G)$ making these cones commute. This morphism is denoted by $\mathrm{Nat}(\alpha, G)$.

$$\begin{tikzcd}
& \mathrm{Nat}(F,G) \arrow[d, dashed, "\mathrm{Nat}(\alpha,G)"] \arrow[ddl, "\gamma_{x,B}"'] \arrow[ddr, "\gamma_{F'(g)(x),B'}"] & \\
& \mathrm{Nat}(F',G) \arrow[dl, "\kappa_{x,B}"'] \arrow[dr, "\kappa_{F'(g)(x),B'}"] & \\
G(B) \arrow[rr, "G(g)"'] & & G(B')
\end{tikzcd}$$

From the above definitions, it follows routinely from uniqueness conditions of the commuting maps into the limits that $\mathrm{Nat}(1_F, G)$ is the identity on $\mathrm{Nat}(F, G)$, and also for $\beta : F'' \Rightarrow F'$ and $\alpha : F' \Rightarrow F$ that $\mathrm{Nat}(\beta, G) \circ \mathrm{Nat}(\alpha, G) = \mathrm{Nat}(\alpha\beta, G) : \mathrm{Nat}(F, G) \to \mathrm{Nat}(F'', G)$.

Hence $\mathrm{Nat}(-, G)$ indeed defines a functor.

8.3.2 Nat Functor as End

Given functors $F : \mathbb{B} \to \mathbf{Set}$ and $G : \mathbb{B} \to \mathbb{A}$, the bifunctor

$$\prod_{F(-)} G(-) : \mathbb{B}^{\mathrm{op}} \times \mathbb{B} \to \mathbb{A}$$

maps an object (B, C) to the product $\prod_{F(B)} G(C)$.

A morphism $(f, g) : (B, C) \to (B', C')$, where $f : B' \to B$ and $g : C \to C'$ is mapped to the morphism

$$\begin{tikzcd}
\prod_{F(B)} G(C) \arrow[r, "\langle G(g)\pi_{F(f)(x)}\rangle_{x\in F(B)}"] \arrow[d, "\pi_{F(f)(x)}"] & \prod_{F(B')} G(C') \arrow[d, "\pi_x"] \\
G(C) \arrow[r, "G(g)"] & G(C')
\end{tikzcd}$$

Where $\langle G(g)\pi_{F(f)(x)}\rangle_{x\in F(B)}$ is the morphism into the product induced by the family of morphisms $G(g)\pi_{F(f)(x)}$ into its components.

The Nat functor can be described as an end of the (bi)functor $\prod_{F(-)} G(-)$,

$$\mathrm{Nat}(F,G) = \int_{B\in\mathbb{B}} \prod_{F(B)} G(B)$$

with family of morphisms

$$\left\langle \chi_B : \mathrm{Nat}(F,G) \to \prod_{F(B)} G(B) \right\rangle_{B\in\mathbb{B}}$$

where for every $B, C \in \mathbb{B}$ and $f : B \to C$ the following diagram commutes:

$$\begin{tikzcd}
 & \prod_{F(B)} G(B) \arrow[dr, "\langle G(f)\pi_x^B\rangle_{x\in F(B)}"] & \\
\mathrm{Nat}(F,G) \arrow[ur, "\chi_B"] \arrow[dr, "\chi_C"'] & & \prod_{F(B)} G(C) \\
 & \prod_{F(C)} G(C) \arrow[ur, "\langle \pi^C_{F(f)(x)}\rangle_{x\in F(B)}"'] &
\end{tikzcd}$$

Where

$$\pi_x^B : \prod_{F(B)} G(B) \to G(B)$$
$$\pi^C_{F(f)(x)} : \prod_{F(C)} G(C) \to G(C)$$

are the projection maps onto the x-th and $F(f)(x)$-th components of the relevant products respectively, and where

$$\langle G(f)\pi_x^B\rangle_{x\in F(B)} : \prod_{F(B)} G(B) \to \prod_{F(B)} G(C)$$
$$\langle \pi^C_{F(f)(x)}\rangle_{x\in F(B)} : \prod_{F(C)} G(C) \to \prod_{F(B)} G(C)$$

are the morphisms into the product $\prod_{F(B)} G(C)$ respectively induced by the families of morphisms $G(f)\pi_x^B$ into its x-th component $G(C)$ and $\pi^C_{F(f)(x)}$ into its x-th component $G(C)$.

$$\begin{tikzcd}
\prod\limits_{F(B)} G(B) \arrow[rr, "\langle G(f)\pi_x^B\rangle_{x\in F(B)}"] \arrow[d, "\pi_x^B"] & & \prod\limits_{F(B)} G(C) \arrow[d, "\pi_x"] \\
G(B) \arrow[rr, "G(f)"] & & G(C)
\end{tikzcd}$$

$$\begin{tikzcd}
\prod\limits_{F(C)} G(C) \arrow[rr, "\langle \pi^C_{F(f)(x)}\rangle_{x\in F(B)}"] \arrow[drr, "\pi^C_{F(f)(x)}"'] & & \prod\limits_{F(B)} G(C) \arrow[d, "\pi_x"] \\
 & & G(C)
\end{tikzcd}$$

What is more, it is universal with this property, in the sense that for any other pair consisting of an object W and family of morphisms

$$\left\langle \omega_B : W \to \prod_{F(B)} G(B) \right\rangle_{B\in\mathbb{B}}$$

with the relevant commutative properties, there is a unique morphism

$$h : W \to \mathrm{Nat}(F, G)$$

such that

$$\chi_B \circ h = \omega_B.$$

See the following diagram:

$$\begin{tikzcd}
 & & \prod\limits_{F(B)} G(B) \arrow[dr, "\langle G(f)\pi_x^B\rangle_{x\in F(B)}"] & \\
W \arrow[r, dashed, "h"] \arrow[urr, bend left, "\omega_B"] \arrow[drr, bend right, "\omega_C"'] & \mathrm{Nat}(F,G) \arrow[ur, "\chi_B"] \arrow[dr, "\chi_C"'] & & \prod\limits_{F(B)} G(C) \\
 & & \prod\limits_{F(C)} G(C) \arrow[ur, "\langle \pi^C_{F(f)(x)}\rangle_{x\in F(B)}"'] &
\end{tikzcd}$$

For each natural transformation $\alpha : F' \to F$, there is also a natural way to obtain $\mathrm{Nat}(\alpha, G) : \mathrm{Nat}(F, G) \to \mathrm{Nat}(F', G)$.

Let

$$\left\langle \nu_B : \mathrm{Nat}(F', G) \to \prod_{F'(B)} G(B) \right\rangle_{B \in \mathbb{B}}$$

be the cone of morphisms corresponding to the end $\mathrm{Nat}(F', G)$.

For any $B \in \mathbb{B}$, there is a morphism $\langle \pi_{\alpha_B} \rangle = \langle \pi^B_{\alpha_B(x)} \rangle_{x \in F'(B)}$ as defined in the diagram below.

$$\begin{tikzcd}
\prod\limits_{F(B)} G(B) \arrow[rr, "\langle \pi^B_{\alpha_B(x)} \rangle_{x \in F'(B)}"] \arrow[drr, "\pi^B_{\alpha_B(x)}"'] & & \prod\limits_{F'(B)} G(B) \arrow[d, "\pi_x"] \\
& & G(B)
\end{tikzcd}$$

It is then routine to check, using the properties of $\mathrm{Nat}(F, G)$ as an end, that the following commutativity condition holds:

$$\langle G(f)\pi^B_x \rangle_{x \in F'(B)} \circ \langle \pi_{\alpha_B} \rangle \circ \chi_B \quad = \quad \langle \pi^C_{F'(f)(x)} \rangle_{x \in F'(B)} \circ \langle \pi_{\alpha_C} \rangle \circ \chi_C$$

$$\begin{tikzcd}
& \prod\limits_{F(B)} G(B) \arrow[r, "\langle \pi_{\alpha_B} \rangle"] & \prod\limits_{F'(B)} G(B) \arrow[dr, "\langle G(f)\pi^B_x \rangle_{x \in F'(B)}"] & \\
\mathrm{Nat}(F, G) \arrow[ur, bend left, "\chi_B"] \arrow[r, "\mathrm{Nat}(\alpha, G)"] \arrow[dr, bend right, "\chi_C"'] & \mathrm{Nat}(F', G) \arrow[ur, "\nu_B"] \arrow[dr, "\nu_C"'] & & \prod\limits_{F'(B)} G(C) \\
& \prod\limits_{F(C)} G(C) \arrow[r, "\langle \pi_{\alpha_C} \rangle"] & \prod\limits_{F'(C)} G(C) \arrow[ur, "\langle \pi^C_{F'(f)(x)} \rangle_{x \in F'(B)}"'] &
\end{tikzcd}$$

Thus, there is a unique morphism

$$\mathrm{Nat}(\alpha, G) : \mathrm{Nat}(F, G) \to \mathrm{Nat}(F', G)$$

with

$$\nu_B \circ \mathrm{Nat}(\alpha, G) = \langle \pi_{\alpha_B} \rangle \circ \chi_B$$

For all $B \in \mathbb{B}$.

That $\mathrm{Nat}(-, G)$ defines a functor follows routinely from uniqueness of the induced map $\mathrm{Nat}(\alpha, G)$.

8.3.3 Equivalence of Limit and End Definitions

Proposition 8.3.3.1. The limit Nat(F,G) as defined in §8.3.1 with the family of morphisms given by

$$\left\langle \left\langle \lambda_{x,B} \right\rangle_{x\in F(B)} : \mathrm{Nat}(F,G) \to \prod_{F(B)} G(B) \right\rangle_{B\in\mathbb{B}}$$

is an end for $\prod_{F(-)} G(-)$, where each

$$\begin{aligned} \lambda_{x,B} :& \mathrm{Nat}(F,G) \to G(B) \\ & \varprojlim GP_F \to GP_F(x,B) \end{aligned}$$

is a component of the limiting cone, and $\langle \lambda_{x,B} \rangle_{x\in F(B)}$ is the induced morphism into the product.

Proof.

$$\begin{tikzcd}
 & & \prod\limits_{F(B)} G(B) \arrow[dr, "\langle G(f)\pi^B_x\rangle_{x\in F(B)}"] & \\
W \arrow[r, dashed, "h"] \arrow[urr, bend left, "\omega_B"] \arrow[drr, bend right, "\omega_C"'] & \mathrm{Nat}(F,G) \arrow[ur, "\langle\lambda_{x,B}\rangle_{x\in F(B)}"] \arrow[dr, "\langle\lambda_{y,C}\rangle_{y\in F(C)}"'] & & \prod\limits_{F(B)} G(C) \\
 & & \prod\limits_{F(C)} G(C) \arrow[ur, "\langle\pi^C_{F(f)(x)}\rangle_{x\in F(B)}"'] &
\end{tikzcd}$$

$$\begin{tikzcd}
 & W \arrow[d, dashed, "h"] \arrow[ddl, "\pi^B_x\omega_B"'] \arrow[ddr, "\pi^C_{F(f)(x)}\omega_C"] & \\
 & \mathrm{Nat}(F,G) \arrow[dl, "\lambda_{x,B}"] \arrow[dr, "\lambda_{F(f)(x),C}"'] & \\
G(B) \arrow[rr, "G(f)"'] & & G(C)
\end{tikzcd}
\qquad
\begin{tikzcd}
\mathrm{Nat}(F,G) \arrow[r, "\langle\lambda_{y,C}\rangle_{y\in F(C)}"] \arrow[d, "\lambda_{y,C}"'] \arrow[dr, "\lambda_{F(f)(x),C}"] & \prod\limits_{F(C)} G(C) \arrow[d, "\pi^C_{F(f)(x)}"] \\
G(C) & G(C)
\end{tikzcd}$$

First, for each $B, C \in \mathbb{B}$ and $f : B \to C$,

$$\begin{aligned}
\langle G(f)\pi_x^B\rangle_{x\in F(B)} \circ \langle \lambda_{x,B}\rangle_{x\in F(B)} &= \langle G(f)\pi_x^B \circ \langle \lambda_{y,B}\rangle_{y\in F(B)}\rangle_{x\in F(B)} \\
&= \langle G(f)\lambda_{x,B}\rangle_{x\in F(B)} \\
&= \langle \lambda_{F(f)(x),C}\rangle_{x\in F(B)} \\
&= \langle \pi^C_{F(f)(x)} \circ \langle \lambda_{y,C}\rangle_{y\in F(C)}\rangle_{x\in F(B)} \\
&= \langle \pi^C_{F(f)(x)}\rangle_{x\in F(B)} \circ \langle \lambda_{y,C}\rangle_{y\in F(C)}
\end{aligned}$$

Now for any object W and family of maps, $(\omega_B : W \to \prod_{F(B)} G(B))_{B\in\mathbb{B}}$ for all $B, C \in \mathbb{B}$ and $f : B \to C$, the following implications hold:

$$\begin{array}{rlcll}
& \langle G(f)\pi_x^B\rangle_{x\in F(B)} \circ \omega_B & = & \langle \pi^C_{F(f)(x)}\rangle_{x\in F(B)} \circ \omega_C & \\
\Rightarrow & \langle G(f)\pi_x^B\omega_B\rangle_{x\in F(B)} & = & \langle \pi^C_{F(f)(x)}\omega_C\rangle_{x\in F(B)} & \\
\Rightarrow & G(f)\pi_x^B\omega_B & = & \pi^C_{F(f)(x)}\omega_C & \text{for all } x \in F(B)
\end{array}$$

Then, the maps $\pi_x^B\omega_B$ form a cone over the diagram of the limit $\mathrm{Nat}(F, G)$ and so there is unique induced map $h : W \to \mathrm{Nat}(F, G)$ such that $\lambda_{x,B}h = \pi_x^B\omega_B$.

Then

$$\langle \lambda_{x,B}\rangle_{x\in F(B)} \circ h = \langle \lambda_{x,B}h\rangle_{x\in F(B)} = \langle \pi_x^B\omega_B\rangle_{x\in F(B)} = \omega_B$$

for all $B \in \mathbb{B}$.

Furthermore, if there were some h' such that $\langle \lambda_{x,B}\rangle_{x\in F(B)} \circ h = \omega_B$ for all $B \in \mathbb{B}$, then also it would satisfy $\lambda_{x,B}h' = \pi_x^B\omega_B$ for all B, and by uniqueness of the induced map into the limit, it would hold that $h = h'$.

Hence $\mathrm{Nat}(F, G)$ is, unique up to isomorphism, the end for $\prod_{F(-)} G(-)$. □

Proposition 8.3.3.2. Let $\alpha : F' \to F$ be a natural transformation.

If the limits $\mathrm{Nat}(F, G)$ and $\mathrm{Nat}(F'G)$ are defined as in §8.3.1, and if the ends are defined as in the previous proposition, then the unique morphism $\mathrm{Nat}(\alpha, G)$ induced between the limits as defined in §8.3.1 is the same morphism induced between the ends.

Proof. Let $\mathrm{Nat}(\alpha, G) : \mathrm{Nat}(F, G) \to \mathrm{Nat}(F', G)$ be the morphism induced between the limits.

Recall that the limit $\mathrm{Nat}(F', G) = \varprojlim GP_{F'}$ has a limiting cone κ with components

$$\kappa_{x,B} : \mathrm{Nat}(F', G) \to GP_{F'}(x, B) = G(B).$$

Then,

$$\begin{aligned}\langle \kappa_{x,B} \rangle_{x\in F'(B)} \circ \mathrm{Nat}(\alpha, G) &= \langle \kappa_{x,B} \circ \mathrm{Nat}(\alpha, G) \rangle_{x\in F'(B)} \\ &= \langle \lambda_{\alpha_B(x),B} \rangle_{x\in F'(B)} \\ &= \langle \pi^B_{\alpha_B(x)} \circ \langle \lambda_{y,B} \rangle_{y\in F(B)} \rangle_{x\in F'(B)} \\ &= \langle \pi^B_{\alpha_B(x)} \rangle_{x\in F'(B)} \circ \langle \lambda_{y,B} \rangle_{y\in F(B)} \\ &= \langle \pi_{\alpha_B} \rangle \circ \langle \lambda_{y,B} \rangle_{y\in F(B)}\end{aligned}$$

Hence

$$\langle \kappa_{x,B} \rangle_{x\in F'(B)} \circ \mathrm{Nat}(\alpha, G) = \langle \pi_{\alpha_B} \rangle \circ \langle \lambda_{y,B} \rangle_{y\in F(B)}$$

and by the uniqueness of the morphism for which this happens, the two definitions of $\mathrm{Nat}(\alpha, G)$ must coincide. □

By the above proof, it is possible to freely switch between the definitions of $\mathrm{Nat}(-, G)$ given as a limit or as an end.

8.4 Adjunction of Hom and Nat

Again, the results of this section (§8.4) are mentioned in [Lei13], but they are worked out in much more detail here.

Theorem 8.4.0.1. $\mathrm{Hom}(-, G)$ is left adjoint to $\mathrm{Nat}(-, G)$ for each $G : \mathbb{B} \to \mathbb{A}$.

$$\begin{tikzcd} A \arrow[rr, bend left, "\mathrm{Hom}(-{,}G)"] \arrow[rr, phantom, "\perp"] & & \left(\mathbf{Set}^{\mathbb{B}}\right)^{\mathrm{op}} \arrow[ll, bend left, "\mathrm{Nat}(-{,}G)"] \end{tikzcd}$$

Proof. For $A \in \mathbb{A}$ and $F : \mathbb{B} \to \mathbf{Set}$, define a map:

$$\phi_{A,F} : \mathrm{Hom}_{\left(\mathbf{Set}^{\mathbb{B}}\right)^{\mathrm{op}}}[\mathrm{Hom}(A,\ G(-)), F] \to \mathrm{Hom}_{\mathbb{A}}[A, \mathrm{Nat}(F, G)]$$

Given a morphism $\beta \in \mathrm{Hom}_{\left(\mathbf{Set}^{\mathbb{B}}\right)^{\mathrm{op}}}[\mathrm{Hom}(A, G(-)), F]$, i.e., a natural transformation $\beta : F \Rightarrow \mathrm{Hom}(A, G(-))$, for each $B \in \mathbb{B}$ there is a map $\beta_B : F(B) \to \mathrm{Hom}(A, G(B))$ which is natural in B.

Let $g : B \to B'$ and $x \in F(B)$. Then $\beta_B(x) : A \to G(B)$, and $\beta_{B'}(F(g)(x)) : A \to G(B')$, and by naturality of β, it holds that $\beta_{B'}(F(g)(x)) = G(g)\beta_B(x)$.

In particular, $\beta_B(x) : A \to GP(x, B)$ forms a cone that commutes with the morphisms $g : (x, B) \to (x', B')$, and so there is a unique morphism

$$q : A \to \varprojlim GP_F = \mathrm{Nat}(F, G)$$

$$\begin{tikzcd}
F(B) \arrow[r, "\beta_B"] \arrow[d, "F(g)"] & \mathrm{Hom}(A, G(B)) \arrow[d, "\mathrm{Hom}(A,G(g))"] \\
F(B') \arrow[r, "\beta_{B'}"] & \mathrm{Hom}(A, G(B'))
\end{tikzcd}
\qquad
\begin{tikzcd}
 & A \arrow[d, dashed, "q"] \arrow[ddl, "\beta_B(x)"'] \arrow[ddr, "\beta_{B'}(F(g)(x))"] & \\
 & \mathrm{Nat}(F, G) \arrow[dl, "\lambda_{x,B}"] \arrow[dr, "\lambda_{F(g)(x),B'}"'] & \\
G(B) \arrow[rr, "G(g)"'] & & G(B')
\end{tikzcd}$$

such that the limiting cone commutes with the cone β.

Define

$$\phi_{A,F}(\beta) = q. \tag{8.4.1}$$

Call the limiting cone $\lambda = (\lambda_{x,B})_{B \in \mathbb{B}, x \in F(B)}$.

If $\beta \neq \gamma$, then there is some B so that $\beta_B \neq \gamma_B$, and so there is some $x \in B$ such that $\beta_B(x) \neq \gamma_B(x)$. For each (x, B), there is a morphism $\lambda_{x,B}$ in the limiting cone. Since there cannot be a single morphism $q : A \to \mathrm{Nat}(A, F)$ such that $\beta_B(x) = \lambda_{x,B} q = \gamma_B(x)$, their induced morphisms q must differ and hence

$$\beta \neq \gamma \qquad \Rightarrow \qquad \phi_{A,F}(\beta) \neq \phi_{A,F}(\gamma)$$

so the map $\phi_{A,F}$ is injective.

For any $q : A \to \mathrm{Nat}(F, G)$, define $\beta_B(x) = \lambda_{x,B} \circ q$. Then for any $g : B \to B'$,

$$\beta_{B'}(F(g)(x)) = (\lambda_{F(g)(x),B'})q = G(g)\lambda_{x,B} \circ q = G(g) \circ \beta_B(x)$$

and so β is a natural transformation. Since q is unique with this property, it is the q given by $q = \phi_{A,F}(\beta)$ and so $\phi_{A,F}$ is surjective.

Hence $\phi_{A,F}$ is a bijection, with inverse given by $\phi^{-1}_{A,F}(q) = \beta$, i.e.,

$$\phi^{-1}_{A,F}(q)_B(x) = \lambda_{x,B} \circ q \tag{8.4.2}$$

$$\phi_{A,F} : \mathrm{Hom}_{(\mathbf{Set}^{\mathbb{B}})^{\mathrm{op}}}[\mathrm{Hom}(A, G(-)), F] \to \mathrm{Hom}_{\mathbb{A}}[A, \mathrm{Nat}(F, G)]$$

Given $f : A' \to A$, there is an induced map $\mathrm{Hom}_{(\mathbf{Set}^{\mathbb{B}})^{\mathrm{op}}}[\mathrm{Hom}(f, G(-)), F]$, also denoted $((f^*)^{\mathrm{op}})^*$.

$((f^*)^{\mathrm{op}})^* : \mathrm{Hom}_{(\mathbf{Set}^{\mathbb{B}})^{\mathrm{op}}}[\mathrm{Hom}(A, G(-)), F] \to \mathrm{Hom}_{(\mathbf{Set}^{\mathbb{B}})^{\mathrm{op}}}[\mathrm{Hom}(A', G(-)), F]$

takes a natural transformation $\beta : F \Rightarrow \mathrm{Hom}(A, G(-))$ and maps it to

$$\mathrm{Hom}(f, G(-)) \circ \beta = (f^*) \circ \beta : F \Rightarrow \mathrm{Hom}(A', G(-))$$

where f^* is right composition by f. Note that $((f^*)^{\mathrm{op}})^*$ is right composition by $(f^*)^{\mathrm{op}}$ in the category $(\mathbf{Set}^{\mathbb{B}})^{\mathrm{op}}$, which corresponds to left composition by f^* with the natural transformation β. For $B \in \mathbb{B}$ and $x \in F(B)$, this is given by

$$((f^*)\beta_B)(x) = (f^*)(\beta_B(x)) = (\beta_B(x)f) : A' \to G(B).$$

$$\begin{tikzcd} & F(B) \arrow[dl, "\beta_B"'] \arrow[dr, "(f^*)\circ\beta_B"] & \\ \mathrm{Hom}(A, GB) \arrow[rr, "f^*"] & & \mathrm{Hom}(A', GB) \end{tikzcd} \qquad \begin{tikzcd} A' \arrow[rr, "f"] \arrow[dr, "\beta_B(x)f"'] & & A \arrow[dl, "\beta_B(x)"] \\ & G(B) & \end{tikzcd}$$

Similarly, there is induced map

$$\mathrm{Hom}_{\mathbb{A}}[f, \mathrm{Nat}(F, G)] = f^* : \mathrm{Hom}_{\mathbb{A}}[A, \mathrm{Nat}(F, G)] \to \mathrm{Hom}_{\mathbb{A}}[A', \mathrm{Nat}(F, G)]$$

which takes $q : A \to \mathrm{Nat}(F, G)$ and maps it to $qf : A' \to \mathrm{Nat}(F, G)$.

$$\begin{tikzcd} A' \arrow[rr, "f"] \arrow[dr, "qf"'] & & A \arrow[dl, "q"] \\ & \mathrm{Nat}(F, G) & \end{tikzcd}$$

It must be shown that $\phi_{A,F}$ is natural in A, i.e., that for and $A, A' \in \mathbb{A}$ and $f : A' \to A$, the following diagram commutes.

$$\begin{tikzcd} \mathrm{Hom}_{(\mathbf{Set}^{\mathbb{B}})^{\mathrm{op}}}[\mathrm{Hom}(A, G(-)), F] \arrow[r, "\phi_{A,F}"] \arrow[d, "((f^*)^{\mathrm{op}})^*"] & \mathrm{Hom}_{\mathbb{A}}[A, \mathrm{Nat}(F, G)] \arrow[d, "f^*"] \\ \mathrm{Hom}_{(\mathbf{Set}^{\mathbb{B}})^{\mathrm{op}}}[\mathrm{Hom}(A', G(-)), F] \arrow[r, "\phi_{A',F}"] & \mathrm{Hom}_{\mathbb{A}}[A', \mathrm{Nat}(F, G)] \end{tikzcd}$$

By inspection of the following diagram and the fact that the induced morphisms into $\mathrm{Nat}(F, G)$ are the unique ones commuting with the relevant cones, it can be seen that

$$\phi_{A',F} \circ \mathrm{Hom}_{(\mathbf{Set}^{\mathbb{B}})^{\mathrm{op}}}[\mathrm{Hom}(f, G(-)), F)] = \mathrm{Hom}(f, \mathrm{Nat}(F, G)) \circ (\phi_{A,F}(\beta))$$

i.e., that

$$\phi_{A',F} \circ ((f^*)^{\mathrm{op}})^* = (f^*) \circ \phi_{A,F}$$

and hence that $\phi_{A,F}$ is natural in A.

$$\begin{tikzcd}
 & & A' \arrow[d, "f"] \arrow[ddll, bend right=40, "\beta_B(x)f"'] \arrow[ddrr, bend left=40, "\beta_{B'}(G(g)(x))f"] \arrow[dd, dashed, bend right=40, "\phi_{A',F}((f^*)\circ\beta)"'] & & \\
 & & A \arrow[d, dashed, "\phi_{A,F}(\beta)"] \arrow[dll, "\beta_B(x)"'] \arrow[drr, "\beta_{B'}(F(g)(x))"] & & \\
 & & \mathrm{Nat}(F,G) \arrow[dll, "\lambda_{x,B}"] \arrow[drr, "\lambda_{F(g)(x),B'}"'] & & \\
G(B) \arrow[rrrr, "G(g)"'] & & & & G(B')
\end{tikzcd}$$

Let F, F' be two functors from $\mathbb{B}$ to **Set**, and let $\alpha : F' \Rightarrow F$.

Note that $\alpha^{\mathrm{op}} : F \to F'$ in $(\mathbf{Set}^{\mathbb{B}})^{\mathrm{op}}$.

Then there is an induced map $\mathrm{Hom}_{(\mathbf{Set}^{\mathbb{B}})^{\mathrm{op}}}[\mathrm{Hom}(A, G(-)), \alpha]$, also denoted $(\alpha^{\mathrm{op}})_*$.

$$(\alpha^{\mathrm{op}})_* : \mathrm{Hom}_{(\mathbf{Set}^{\mathbb{B}})^{\mathrm{op}}}[\mathrm{Hom}(A, G(-)), F] \to \mathrm{Hom}_{(\mathbf{Set}^{\mathbb{B}})^{\mathrm{op}}}[\mathrm{Hom}(A, G(-)), F']$$

corresponds to left composition in $(\mathbf{Set}^{\mathbb{B}})^{\mathrm{op}}$ and hence right composition with a natural transformation $\beta : F \Rightarrow \mathrm{Hom}(A, G(-))$ as

$$(\alpha^*)(\beta) = \beta \circ \alpha : F' \Rightarrow \mathrm{Hom}(A, G(-))$$

There is also a map

$$\mathrm{Hom}_{\mathbb{A}}[A, \mathrm{Nat}(\alpha, G)] = \mathrm{Nat}(\alpha, G)_*$$

$$\begin{tikzcd}
\mathrm{Hom}_{(\mathbf{Set}^{\mathbb{B}})^{\mathrm{op}}}[\mathrm{Hom}(A, G(-)), F] \arrow[r, "\phi_{A,F}"] \arrow[d, "(\alpha^{\mathrm{op}})_*"] & \mathrm{Hom}_{\mathbb{A}}[A, \mathrm{Nat}(F, G)] \arrow[d, "\mathrm{Nat}(\alpha, G)_*"] \\
\mathrm{Hom}_{(\mathbf{Set}^{\mathbb{B}})^{\mathrm{op}}}[\mathrm{Hom}(A, G(-)), F'] \arrow[r, "\phi_{A,F'}"] & \mathrm{Hom}_{\mathbb{A}}[A, \mathrm{Nat}(F', G)]
\end{tikzcd}$$

For the above square to be commutative is to say that for any $\beta : F \Rightarrow \mathrm{Hom}(A, G(-))$, it holds that

$$\phi_{A,F'}(\beta\alpha) = \mathrm{Nat}(\alpha, G) \circ (\phi_{A,F}(\beta)).$$

Remark. Recall that $(\alpha^{\mathrm{op}})_*(\beta^{\mathrm{op}})$ is $\alpha^{\mathrm{op}}\beta^{\mathrm{op}}$ which corresponds to $\beta\alpha$. $\triangleleft$

Recall that $\phi_{A,F'}(\beta\alpha)$ is the unique map q such that, for the cone consisting of maps $(\beta\alpha)_B(x) : A \to G(B)$ for $B \in \mathbb{B}$ and $x \in F'(B)$ and the limit cone consisting of maps $\kappa_{x,B} : \mathrm{Nat}(F', G) \to G(B)$, it holds that $\kappa_{x,B} \circ q = (\beta\alpha)_B(x)$.

Then to show $\mathrm{Nat}(\alpha, G)\phi_{A,F}(\beta) = \phi_{A,F'}(\beta\alpha)$ is to show that the former also has the above composition property. I.e., $\kappa_{x,B} \circ \mathrm{Nat}(\alpha, G) \circ \phi_{A,F}(\beta) = (\beta\alpha)_B(x)$ for all B and $x \in F(B)$.

Now, $\phi_{A,F}(\beta)$ is the unique map with the property that $\lambda_{x,B} \circ \phi_{A,F}(\beta) = \beta_B(x)$.

Using the above property, and the definitions of $\mathrm{Nat}(\alpha, G)$ and γ (defined in Equation 8.3.3 of §8.3.1), it follows that:

$$\begin{aligned} \kappa_{x,B} \circ \mathrm{Nat}(\alpha, G) \circ \phi_{A,F}(\beta) &= \gamma_{x,B} \circ \phi_{A,F}(\beta) \\ &= \lambda_{\alpha_B(x),B} \circ \phi_{A,F}(\beta) \\ &= \beta_B(\alpha_B(x)) \end{aligned}$$

$$\begin{tikzcd}
 & & A \arrow[d, "\phi_{A,F}(\beta)"] \arrow[dd, dashed, bend right=40, "\phi_{A,F'}(\beta\alpha)"'] \arrow[dddll, bend right=40, "\beta_B(\alpha_B(x))"'] \arrow[dddrr, bend left=40, "\beta_{B'}(\alpha_{B'}F'(g)(x))"] & & \\
 & & \mathrm{Nat}(F,G) \arrow[d, dashed, "\mathrm{Nat}(\alpha,G)"] \arrow[ddll, "\gamma_{x,B}"'] \arrow[ddrr, "\gamma_{F'(g)(x),B'}"] & & \\
 & & \mathrm{Nat}(F',G) \arrow[dll, "\kappa_{x,B}"'] \arrow[drr, "\kappa_{F'(g)(x),B'}"] & & \\
G(B) \arrow[rrrr, "G(g)"'] & & & & G(B')
\end{tikzcd}$$

Hence $\phi_{A,F}$ is natural in F.

The above are all the properties necessary for an adjunction. $\square$

8.4.1 Unit and Counit

The unit for this adjunction is given at each $A \in \mathbb{A}$ by

$$\eta_A = \phi_{A,\mathrm{Hom}(A,G(-))}\left(1_{\mathrm{Hom}(A,G(-))}\right)$$
$$\eta_A : A \to \mathrm{Nat}(\mathrm{Hom}(A, G(-)), G).$$

Its co-unit is given at each $F : \mathbb{B} \to \mathbf{Set}$ by

$$(\epsilon_F)^{\mathrm{op}} = (\phi_{\mathrm{Nat}(F,G),F})^{-1}\left(1_{\mathrm{Nat}(F,G)}\right)$$
$$(\epsilon_F)^{\mathrm{op}} : \mathrm{Hom}(\mathrm{Nat}(F, G), G(-)) \to F$$

where ϵ_F is then a natural transformation

$$\epsilon_F : F \Rightarrow \mathrm{Hom}(\mathrm{Nat}(F, G), G(-)).$$

Recall from Lemma 8.2.1.1 that the functor G is codense if and only if the left adjoint $\mathrm{Hom}(-, G)$ is full and faithful.

This happens if and only if the unit η_A is an isomorphism at each $A \in \mathbb{A}$. In this case each $A \in \mathbb{A}$ is a limit of images under $\mathrm{Hom}(-, G)$ of functors F.

8.5 Adjunction-Induced Monad

The adjunction induces a monad

$$\mathbf{T}^G = (T^G, \eta^G, \mu^G)$$

where T^G is given by the composition

$$T^G = \mathrm{Nat}(\mathrm{Hom}(-, G), G),$$

$\eta^G = \eta$ is the unit of the adjunction and the multiplication is given as

$$\mu^G = \mathrm{Nat}(-, G)\,\epsilon\,\mathrm{Hom}(-, G)$$

where ϵ is the co-unit of the adjunction.

At an object $A \in \mathbb{A}$

$$T^G(A) = \mathrm{Nat}(\mathrm{Hom}(A, G(-)), G)$$

$$\eta^G_A = \eta_A : A \to T(A)$$
$$\eta^G_A : A \to \mathrm{Nat}(\mathrm{Hom}(A, G(-)), G)$$

$$\mu^G_A : T^{\mathbf{2}}(A) \to T(A)$$
$$\mu^G_A : \mathrm{Nat}(\mathrm{Hom}(\mathrm{Nat}(\mathrm{Hom}(A, G(-)), G), G), G) \to \mathrm{Nat}(\mathrm{Hom}(A, G(-)), G)$$
$$\mu^G_A = \mathrm{Nat}((\epsilon_{\mathrm{Hom}(A,G(-))})^{\mathrm{op}}, G)$$

8.6 Adjunction for Inclusion of FinSet into Set

This section (§8.6) still follows what is mentioned in [Lei13], and again are worked out in detail here.

Let $\mathbb{B} = \mathbf{FinSet}$, $\mathbb{A} = \mathbf{Set}$ and $G : \mathbf{FinSet} \to \mathbf{Set}$ be the inclusion of the category of finite sets into the category of sets.

Then there are functors $\mathrm{Hom}(-, G)$ and $\mathrm{Nat}(-, G)$ and adjunction

$$\begin{tikzcd} \mathbf{Set} \arrow[rr, bend left, "\mathrm{Hom}(-,G)"] \arrow[rr, phantom, "\perp"] & & \left(\mathbf{Set}^{\mathbf{FinSet}}\right)^{\mathrm{op}} \arrow[ll, bend left, "\mathrm{Nat}(-,G)"] \end{tikzcd}$$

For $B \in \mathbf{Set}$ its image $\mathrm{Hom}(B, G(-))$ is the functor that maps a finite set $A \in \mathbf{FinSet}$ to the set of morphisms $\mathrm{Hom}(B, G(A))$ from B to $G(A)$.

For a functor $F : \mathbf{FinSet} \to \mathbf{Set}$, its image $\mathrm{Nat}(F, G)$ is the end of $\prod_{F(-)} G(*)$,

$$\mathrm{Nat}(F, G) = \int_{B\in\mathbb{B}} \prod_{F(B)} G(B)$$

with family of morphisms

$$\left\langle \chi_B : \mathrm{Nat}(F, G) \to \prod_{F(B)} G(B) \right\rangle_{B\in\mathbb{B}}$$

where for every $B, C \in \mathbb{B}$ and $f : B \to C$ the following diagram commutes:

$$\begin{tikzcd} & \prod_{F(B)} G(B) \arrow[dr, "\langle G(f)\pi_x^B\rangle_{x\in F(B)}"] & \\ \mathrm{Nat}(F,G) \arrow[ur, "\chi_B"] \arrow[dr, "\chi_C"'] & & \prod_{F(B)} G(C) \\ & \prod_{F(C)} G(C) \arrow[ur, "\langle \pi^C_{F(f)(x)}\rangle_{x\in F(B)}"'] & \end{tikzcd} \tag{8.6.1}$$

It is claimed that this end is given by the set of natural transformations from F to G and for each $B \in \mathbb{B}$ the morphism χ_B takes natural transformation $\alpha : F \Rightarrow G$ to family $\langle \alpha_B(x)\rangle_{x\in F(B)}$. This is now proved.

If $\mathrm{Nat}(F, G)$ were written as a limit, with limit cone given by $\lambda_{x,B}$, then since $\chi_B = \langle \lambda_{x,B} \rangle_{x \in F(B)}$, then $\lambda_{x,B}$ can be obtained by

$$\lambda_{x,B} = \pi_x^B \circ \chi_B. \tag{8.6.2}$$

This means for $\alpha \in \mathrm{Nat}(F, G)$, the morphism $\lambda_{x,B}$ should take the natural transformation α to $\alpha_B(x)$.

In other words, $\lambda_{x,B}(\alpha)$ should be the evaluation of α_B at $x \in F(B)$ and χ_B the family of such evaluations at each $x \in F(B)$. Now define $\mathrm{Nat}(F, G)$ as the set of natural transformations from F to G and define χ_B as above.

Lemma 8.6.0.1. The above set $\mathrm{Nat}(F, G)$ and family of morphisms χ_B is commutative with respect to the Diagram 8.6.1.

Proof.

$$\begin{aligned}
\langle G(f)\pi_x^B \rangle_{x \in F(B)} \langle \alpha_B(y) \rangle_{y \in F(B)} &= \langle G(f)\pi_x^B \langle \alpha_B(y) \rangle_{y \in F(B)} \rangle_{x \in F(B)} \\
&= \langle G(f)\alpha_B(x) \rangle_{x \in F(B)} \\
&= \langle \alpha_C(F(f)(x)) \rangle_{x \in F(B)} \\
&= \langle \pi^C_{F(f)(x)} \langle \alpha_C(y) \rangle_{y \in F(C)} \rangle_{x \in F(B)} \\
&= \langle \pi^C_{F(f)(x)} \rangle_{x \in F(B)} \langle \alpha_C(y) \rangle_{y \in F(C)}
\end{aligned}$$

□

Next, it is shown that it is universal with this property. For any other pair consisting of an object W and family of morphisms

$$\left\langle \omega_B : W \to \prod_{F(B)} G(B) \right\rangle_{B \in \mathbb{B}}$$

with the relevant commutative properties, it must be shown that there is a unique morphism

$$h : W \to \mathrm{Nat}(F, G)$$

such that

$$\chi_B \circ h = \omega_B.$$

See the following diagram:

$$
\begin{tikzcd}
 & & \prod\limits_{F(B)} G(B) \arrow[dr, "\langle G(f)\pi_x^B\rangle_{x\in F(B)}"] & \\
W \arrow[r, dashed, "h"] \arrow[urr, bend left, "\omega_B"] \arrow[drr, bend right, "\omega_C"'] & \mathrm{Nat}(F,G) \arrow[ur, "\chi_B"] \arrow[dr, "\chi_C"] & & \prod\limits_{F(B)} G(C) \\
 & & \prod\limits_{F(C)} G(C) \arrow[ur, "\langle \pi^C_{F(f)(x)}\rangle_{x\in F(B)}"'] &
\end{tikzcd}
$$

Lemma 8.6.0.2. If such an h exists, it is unique with this property.

Proof. For every $B \in \mathbb{B}$ and $w \in W$, it must hold that

$$
\begin{aligned}
\chi_B \circ h(w) &= \langle h(w)_B(x)\rangle_{x\in F(B)} \\
&= \omega_B(w) \\
&= \langle \pi_x^B \circ \omega_B(w)\rangle_{x\in F(B)}
\end{aligned}
$$

hence

$$\langle \pi_x^B \circ \omega_B(w)\rangle_{x\in F(B)} = \langle h(w)_B(x)\rangle_{x\in F(B)}$$

and so, for every $x \in F(B)$

$$h(w)_B(x) = \pi_x^B \circ \omega_B(w)$$

and since each $h(w)$ is a natural transformation, then this uniquely defines h. $\square$

Lemma 8.6.0.3. If $h(w)$ is defined as in the above proof, then it is a natural transformation.

Proof. Let $f : B \to C$. Then since

$$\langle G(f)\pi_x^B\rangle_{x\in F(B)} \circ \omega_B = \langle \pi^C_{F(f)(x)}\rangle_{x\in F(B)} \circ \omega_C$$

it holds for every $x \in F(B)$ that

$$G(f)(\pi_x^B \circ \omega_B) = \pi^C_{F(f)(x)} \circ \omega_C$$

and so

$$
\begin{aligned}
h(w)_C(F(f)(x)) &= \pi^C_{F(f)(x)} \circ \omega_C(w) \\
&= G(f)(\pi_x^B \circ \omega_B(w)) \\
&= G(f)h(w)_B(x)
\end{aligned}
$$

$\square$

That is, for any $w \in W$, its image under h is a natural transformation from F to G given at object $B \in \mathbb{B}$ by

$$h(w)_B : F(B) \to G(B) : x \mapsto \pi_x^B \circ \omega_B(w)$$

and so, also, $\chi_B \circ h = \omega_B$ for each B, and so the commutativity conditions for h are satisfied.

For $\alpha : F' \Rightarrow F$, the morphism $\mathrm{Nat}(\alpha, B) : \mathrm{Nat}(F, G) \to \mathrm{Nat}(F', G)$ is given by right-composition with α. In other words, for $\beta : F \Rightarrow G$, its image $\mathrm{Nat}(\alpha, G)(\beta)$ is $\beta\alpha$.

Hence, as a functor from $\left(\mathbf{Set}^{\mathbf{FinSet}}\right)^{\mathrm{op}}$ to $\mathbf{Set}$, the functor $\mathrm{Nat}(-, G)$ acts the same as the (contravariant) Hom functor $\mathrm{Hom}_{\left(\mathbf{Set}^{\mathbf{FinSet}}\right)^{\mathrm{op}}}(-, G)$.

8.6.1 Adjunction

The natural bijection

$$\phi_{A,F} : \mathrm{Hom}_{\left(\mathbf{Set}^{\mathbf{FinSet}}\right)^{\mathrm{op}}}[\mathrm{Hom}(A, G(-)), F] \to \mathrm{Hom}_{\mathbf{Set}}[A, \mathrm{Nat}(F, G)]$$

is given as follows:

For $\beta : F \to \mathrm{Hom}(A, G(-))$, its image is $\phi_{A,F}(\beta) : A \to \mathrm{Nat}(F, G)$, which by Equation 8.4.1 will be the unique morphism into the limit induced by $\beta_B(x) : A \to G(B)$.

From the proof of Proposition 8.3.3.1, this is the same as the unique morphism induced by defining $\omega_B : A \to \prod_{F(B)} G(B)$ as $\omega_B(a) = \langle \beta_B(x)(a) \rangle_{x \in F(B)}$, for each $a \in A$.

Then the image of $a \in A$ under the induced morphism $h : A \to \mathrm{Nat}(F, G)$ is a natural transformation given at $B \in \mathbb{B}$ by

$$\begin{aligned} h(a)_B(x) &= \pi_x^B \omega_B(a) \\ &= \pi_x^B \left(\langle \beta_B(y)(a) \rangle_{y \in F(B)} \right) \\ &= \beta_B(x)(a). \end{aligned}$$

So at each object $A \in \mathbf{Set}$, functor $F : \mathbf{FinSet} \to \mathbf{Set}$, natural transformation $\beta : F \Rightarrow \mathrm{Hom}(A, G(-))$, object $B \in \mathbf{FinSet}$, element $a \in A$, and element $x \in F(B)$,

$$\begin{aligned} &\phi_{A,F}(\beta) : && A \to \mathrm{Nat}(F, G) \\ &\phi_{A,F}(\beta)(a) : && F \Rightarrow G \\ &(\phi_{A,F}(\beta)(a))_B : && F(B) \to G(B) \\ & && x \mapsto \beta_B(x)(a). \end{aligned}$$

In full

$$(\phi_{A,F}(\beta)(a))_B(x) = \beta_B(x)(a).$$

The inverse $\phi^{-1}_{A,F}$ is given as follows:

Given object $A \in \mathbf{Set}$, functor $F : \mathbf{FinSet} \to \mathbf{Set}$, object $B \in \mathbf{FinSet}$ and morphism $q : A \to \mathrm{Nat}(F, G)$. For $a \in A$ and $x \in F(B)$,

$$\begin{aligned}
\phi^{-1}_{A,F}(q) : &\quad F \Rightarrow \mathrm{Hom}(A, G(-)) \\
(\phi^{-1}_{A,F}(q))_B : &\quad F(B) \to \mathrm{Hom}_{\mathbf{Set}}(A, G(B)) \\
(\phi^{-1}_{A,F}(q))_B(x) : &\quad A \to G(B) \\
&\quad a \mapsto q(a)_B(x)
\end{aligned}$$

Again in full:

$$((\phi^{-1}_{A,F}(q))_B(x))(a) = q(a)_B(x)$$

This follows from Equation 8.4.2 in §8.4, and Equation 8.6.2 in §8.6.

8.6.2 Unit of the Adjunction

The unit for this adjunction is given at each $A \in \mathbf{Set}$ by

$$\eta_A = \phi_{A,\mathrm{Hom}(A,G(-))}\left(1_{\mathrm{Hom}(A,G(-))}\right)$$

$$\eta_A : A \to \mathrm{Nat}(\mathrm{Hom}(A, G(-)), G).$$

Where at each element $a \in A$, object $B \in \mathbf{FinSet}$, and morphism $f : A \to G(B)$,

$$\begin{aligned}
\phi_{A,\mathrm{Hom}(A,G(-))}\left(1_{\mathrm{Hom}(A,G(-))}\right) : &\quad A \to \mathrm{Nat}(\mathrm{Hom}(A, G(-)), G) \\
\phi_{A,\mathrm{Hom}(A,G(-))}\left(1_{\mathrm{Hom}(A,G(-))}\right)(a) : &\quad \mathrm{Hom}(A, G(-)) \Rightarrow G \\
\left(\phi_{A,\mathrm{Hom}(A,G(-))}\left(1_{\mathrm{Hom}(A,G(-))}\right)(a)\right)_B : &\quad \mathrm{Hom}(A, G(B)) \to G(B) \\
&\quad (f : A \to G(B)) \mapsto 1(f)(a) = f(a).
\end{aligned}$$

In full

$$(\phi_{A,\mathrm{Hom}(A,G(-))}(1_{\mathrm{Hom}(A,G(-))})(a))_B(f) = 1_{\mathrm{Hom}(A,G(-))}(f)(a) = f(a).$$

More compactly

$$\eta(a)_B(f) = f(a).$$

Hence the image of an element $a \in A$ of the unit $\eta : A \to \mathrm{Nat}(\mathrm{Hom}(A, G(-)), G)$ is the 'evaluation at a' natural transformation from $\mathrm{Hom}(A, G(-))$ to G.

The functor G is codense if and only if η_A is an isomorphism at each $A \in \mathbb{A}$. This does not happen in this case, since there are many more natural transformations than just the evaluations at each element. Hence the inclusion $G : \mathbf{FinSet} \to \mathbf{Set}$ is not codense.

8.6.3 Co-Unit of the Adjunction

The co-unit is given at each $F : \mathbf{FinSet} \to \mathbf{Set}$ by

$$(\epsilon_F)^{\mathrm{op}} = (\phi_{\mathrm{Nat}(F,G),F})^{-1}\left(1_{\mathrm{Nat}(F,G)}\right)$$

$$(\epsilon_F)^{\mathrm{op}} : \mathrm{Hom}(\mathrm{Nat}(F, G), G(-)) \to F$$

where ϵ_F is then a natural transformation

$$\epsilon_F : F \Rightarrow \mathrm{Hom}(\mathrm{Nat}(F, G), G(-)).$$

By Equation 8.4.2 in §8.4, this is the natural transformation given by $(\epsilon_F)_B(x) = \lambda_{x,B}$, where λ is the limiting cone given by

$$\mathrm{Nat}(F, G) = \varprojlim GP_F$$

By Equation 8.6.2 in §8.6, recall that $\lambda_{x,B}(\alpha)$ is the evaluation of α_B at $x \in F(B)$, for $\alpha \in \mathrm{Nat}(F, G)$.

Hence

$$(\epsilon_F)_B : F(B) \Rightarrow \mathrm{Hom}(\mathrm{Nat}(F, G), G(B)).$$

takes $x \in F(B)$ to the evaluation-at-x morphism from $\mathrm{Nat}(F, G)$ to $G(B)$ which itself takes a natural transformation $\alpha \in \mathrm{Nat}(F, G)$ to $\alpha_B(x)$

8.7 Codensity Monad of G : FinSet $\to$ Set

This section (§8.7) continues to work out in detail the results in [Lei13].

The adjunction induces a monad

$$\mathbf{T}^G = (T^G, \eta^G, \mu^G)$$

where T^G is given by the composition

$$T^G = \mathrm{Nat}(\mathrm{Hom}(-, G), G),$$

$\eta^G = \eta$ is the unit of the adjunction and the multiplication is given as

$$\mu^G = \mathrm{Nat}(-, G)\ \epsilon^{\mathrm{op}}\ \mathrm{Hom}(-, G)$$

where ϵ is the co-unit of the adjunction.

This is the codensity monad of the functor G.

At an object $A \in \mathbb{A}$

$$T^G(A) = \mathrm{Nat}(\mathrm{Hom}(A, G), G)$$

$$\eta_A^G = \eta_A : A \to T(A)$$
$$\eta_A^G : A \to \mathrm{Nat}(\mathrm{Hom}(A, G(-)), G)$$

$$\mu_A^G : T^{\mathbf{2}}(A) \to T(A)$$
$$\mu_A^G : \mathrm{Nat}[\mathrm{Hom}(\mathrm{Nat}[\mathrm{Hom}(A, G(-)), G], G(-)), G] \to \mathrm{Nat}(\mathrm{Hom}(A, G(-)), G)$$
$$\mu_A^G = \mathrm{Nat}((\epsilon_{\mathrm{Hom}(A,G(-))})^{\mathrm{op}}, G)$$

8.8 Codensity Monad is Ultrafilter

The following section follows [Lei13] very closely.

For set X finite set B, and each ultrafilter $\mathcal{U}$ on $\mathcal{P}(X)$, a set map $f : X \to B$ gives rise to a partition of X into finitely many elements.

Define an 'integration' operator:

$$\int_X^B - \,\mathrm{d}\mathcal{U} : \mathrm{Hom}_{\mathbf{Set}}(X, G(B)) \to B$$

Which takes a set map $f : X \to B$ to the unique $b \in B$ with $f^{-1}(b) \in \mathcal{U}$.

Hence

$$f^{-1}\left(\int_X^B f \,\mathrm{d}\mathcal{U}\right) \in \mathcal{U}$$

and is the unique element of B doing so.

For B a singleton, the existence of such a unique element is obvious. If $|B| = 2$, then it follows directly from the definition of an ultrafilter, and if $|B| > 2$, then it follows by the partition condition of §3.4.

For each $X \in \mathbf{Set}$ there is a functor

$$\mathrm{Hom}_{\mathbf{Set}}(X, G(-)) : \mathbf{FinSet} \to \mathbf{Set}.$$

Given that $G : \mathbf{FinSet} \to \mathbf{Set}$ is the inclusion functor, there is a natural transformation from $\mathrm{Hom}_{\mathbf{Set}}(X, G(-))$ to G.

Lemma 8.8.0.1. There is natural transformation

$$\int_X^{-} \mathrm{d}\mathcal{U} : \mathrm{Hom}_{\mathbf{Set}}(X, G(-)) \Rightarrow G$$

given at each $B \in \mathbf{FinSet}$ by

$$\left(\int_X^{-} \mathrm{d}\mathcal{U}\right)_B = \int_X^B - \,\mathrm{d}\mathcal{U}.$$

Proof.

$$\begin{tikzcd}
\mathrm{Hom}_{\mathbf{Set}}(X, B) \arrow[r, "\int_X^B - \,\mathrm{d}\mathcal{U}"] \arrow[d, "\theta_*"] & B \arrow[d, "\theta"] \\
\mathrm{Hom}_{\mathbf{Set}}(X, B') \arrow[r, "\int_X^{B'} - \,\mathrm{d}\mathcal{U}"] & B'
\end{tikzcd}$$

Let $\theta : B \to B'$. Then it must be shown for any $f : X \to B$ that

$$\theta\left(\int_X^B f \,\mathrm{d}\mathcal{U}\right) = \int_X^{B'} \theta f \,\mathrm{d}\mathcal{U}$$

This follows since

$$\begin{aligned}
(\theta f)^{-1}\left(\theta\left(\int_X^B f \,\mathrm{d}\mathcal{U}\right)\right) &= f^{-1}\left(\theta^{-1}\left(\theta\left(\int_X^B f \,\mathrm{d}\mathcal{U}\right)\right)\right) \\
&\supseteq f^{-1}\left(\int_X^B f \,\mathrm{d}\mathcal{U}\right) \in \mathcal{U}.
\end{aligned}$$

Hence, since supersets of ultrafilter elements must be in the ultrafilter.

$$(\theta f)^{-1}\left(\theta\left(\int_X^B f \,\mathrm{d}\mathcal{U}\right)\right) \in \mathcal{U}$$

and since $\int_X^B \theta f \,\mathrm{d}\mathcal{U}$ is unique with this property, they must be equal. □

Lemma 8.8.0.2. For any sets X and Y, map $p : X \to Y$, and ultrafilter $\mathcal{U} \in U(X)$, the diagram

$$\begin{tikzcd}
\mathrm{Hom}_{\mathbf{Set}}(Y, G(-)) \arrow[rr, "\mathrm{Hom}(p, G(-))"] \arrow[dr, "\int_Y^{-} \mathrm{d}(p_*\mathcal{U})"'] & & \mathrm{Hom}_{\mathbf{Set}}(X, G(-)) \arrow[dl, "\int_X^{-} \mathrm{d}\mathcal{U}"] \\
& G &
\end{tikzcd}$$

commutes. I.e.,

$$\left(\int_X^- \mathrm{d}\mathcal{U}\right) \circ \mathrm{Hom}(p, G(-)) \qquad = \qquad \left(\int_Y^- \mathrm{d}(p_*\mathcal{U})\right)$$

Where $p_*\mathcal{U}$ is defined as in §3.5.1.

Proof. For any finite set $B \in \mathbf{FinSet}$ and map $f : Y \to B$,

$$\left(\left(\int_X^B - \mathrm{d}\mathcal{U}\right) \circ \mathrm{Hom}(p, G(B))\right)(f) = \left(\int_X^B - \mathrm{d}\mathcal{U}\right)(fp) = \int_X^B fp \,\mathrm{d}\mathcal{U}.$$

Now

$$(fp)^{-1}\left(\int_X^B f\mathrm{d}(p_*\mathcal{U})\right) = p^{-1}\left(f^{-1}\left(\int_Y^B f \,\mathrm{d}(p_*\mathcal{U})\right)\right)$$

and,

$$f^{-1}\left(\int_Y^B f \,\mathrm{d}(p_*\mathcal{U})\right) \in p_*\mathcal{U}$$

so by definition of $p_*\mathcal{U}$ from §3.5.1, this means

$$p^{-1}\left(f^{-1}\left(\int_Y^B f \,\mathrm{d}(p_*\mathcal{U})\right)\right) \in \mathcal{U}.$$

Now, since $\int_X^B fp \,\mathrm{d}(\mathcal{U}))$ is unique such that its inverse image under fp is in $\mathcal{U}$, then

$$\left(\left(\int_X^B - \mathrm{d}\mathcal{U}\right) \circ \mathrm{Hom}(p, G(B))\right)(f) = \int_X^B fp \,\mathrm{d}\mathcal{U} = \int_Y^B f \,\mathrm{d}(p_*\mathcal{U})$$

and so the two maps are equal for any finite $B \in \mathbf{FinSet}$ and so the natural transformations of the lemma's statement are equal. □

Corollary 8.8.0.3. Given the ultrafilter functor U and functor of the codensity monad T^G of the inclusion $G : \mathbf{FinSet} \hookrightarrow \mathbf{Set}$, there is a natural transformation

$$\int_- \mathrm{d} : U \Rightarrow T^G$$

with component at each $X \in \mathbf{Set}$ given by

$$\left(\int_- \mathrm{d}\right)_X = \int_X \mathrm{d}(-)$$

where

$$\int_X \mathrm{d}(-) : \qquad U(X) \to T^G(X) : \qquad \mathcal{U} \mapsto \int_X^- \mathrm{d}\mathcal{U}.$$

Proof. Recall that $T^G(X) = \mathrm{Nat}(\mathrm{Hom}(X, G(-)), G)$ and by lemma 8.8.0.1 that

$$\int_X^- \mathrm{d}\mathcal{U} : \mathrm{Hom}_{\mathbf{Set}}(X, G(-)) \Rightarrow G$$

is a natural transformation where the second component must be a finite set, so that the map is well-defined at each $B \in \mathbf{FinSet}$.

For this to be a natural transformation means that

$$T^G(p) \circ \int_X \mathrm{d}(-) = \int_Y \mathrm{d}(-) \circ U(p)$$

for any X, Y and $p : X \to Y$. However, at a particular $\mathcal{U} \in U(X)$,

$$\begin{aligned}\left(T^G(p) \circ \int_X^- \mathrm{d}(-)\right)(\mathcal{U}) &= T^G(p)\left(\int_X^- \mathrm{d}(\mathcal{U})\right)\\ &= \mathrm{Nat}(\mathrm{Hom}(p, G), G)\left(\int_X^- \mathrm{d}(\mathcal{U})\right)\\ &= \left(\int_X^- \mathrm{d}(\mathcal{U})\right) \circ \mathrm{Hom}(p, G).\end{aligned}$$

Then by lemma 8.8.0.2

$$\begin{aligned}\int_X^- \mathrm{d}\mathcal{U} \circ \mathrm{Hom}(p, G) &= \int_Y^- \mathrm{d}(p_*\mathcal{U})\\ &= \left(\int_Y \mathrm{d}(-)\right)(p_*\mathcal{U})\\ &= \left(\int_Y \mathrm{d}(-)\right)(U(p)(\mathcal{U}))\\ &= \left(\int_Y \mathrm{d}(-) \circ U(p)\right)(\mathcal{U})\end{aligned}$$

□

Lemma 8.8.0.4. At each set $X \in \mathbf{Set}$, the map

$$\int_X \mathrm{d}(-) : U(X) \to T^G(X)$$

is injective.

Proof. It is shown for any $\mathcal{U}, \mathcal{V} \in U(X)$ that if $\int \mathrm{d}\mathcal{U} = \int \mathrm{d}\mathcal{V}$ then $\mathcal{U} = \mathcal{V}$.

Consider the set $\mathbf{2} = \{0, 1\}$ and for each $Y \subseteq X$ define $\chi_Y : X \to \mathbf{2}$ by

$$\chi_Y(x) = \begin{cases} 1, & x \in Y \\ 0, & \text{otherwise.} \end{cases} \tag{8.8.1}$$

The inverses are given by $\chi_Y^{-1}(1) = Y$ and $\chi_Y^{-1}(0) = X - Y$. Recall for any ultrafilter $\mathcal{U}$ on $\mathcal{P}(X)$, that exactly one of Y and $(X - Y)$ is in $\mathcal{U}$. Then exactly one of $\chi_Y^{-1}(1)$ or $\chi_Y^{-1}(0)$ is in $\mathcal{U}$.

Thus, since $\int_X^{\mathbf{2}} \chi_Y \,\mathrm{d}\mathcal{U}$ denotes the unique element of $\mathbf{2}$ such that

$$\chi_Y^{-1}\left(\int_X^{\mathbf{2}} \chi_Y \,\mathrm{d}\mathcal{U}\right) \in \mathcal{U}$$

then

$$\int_X^{\mathbf{2}} \chi_Y \,\mathrm{d}\mathcal{U} = \begin{cases} 1, & Y \in \mathcal{U} \\ 0, & \text{otherwise} \end{cases}$$

and so, given $\int_X^{-} \mathrm{d}\mathcal{U}$, one can obtain $\mathcal{U}$ by

$$\mathcal{U} = \left\{ Y \subseteq X : \int_X^{\mathbf{2}} \chi_Y \,\mathrm{d}\mathcal{U} = 1 \right\}.$$

Assume $\int \mathrm{d}\mathcal{U} = \int \mathrm{d}\mathcal{V}$. Then $\int_X^{\mathbf{2}} \chi_Y \,\mathrm{d}\mathcal{U} = \int_X^{\mathbf{2}} \chi_Y \,\mathrm{d}\mathcal{V}$, so

$$\mathcal{U} = \left\{ Y \subseteq X : \int_X^{\mathbf{2}} \chi_Y \,\mathrm{d}\mathcal{U} = 1 \right\} = \left\{ Y \subseteq X : \int_X^{\mathbf{2}} \chi_Y \,\mathrm{d}\mathcal{V} = 1 \right\} = \mathcal{V}.$$

□

Lemma 8.8.0.5. At each set $X \in \mathbf{Set}$, the map

$$\int_X \mathrm{d}(-) : U(X) \to T^G(X)$$

is surjective.

Proof. Let $\alpha \in T^G(X)$, i.e., $\alpha : \mathrm{Hom}_{\mathbf{Set}}(X, G(-)) \Rightarrow G$.

It is shown there is some $\mathcal{U}_\alpha \in U(X)$ such that $\int_X \mathrm{d}\mathcal{U}_\alpha = \alpha$.

Let $B \in \mathbf{FinSet}$ be a finite set and $f : X \to B$ be a map (which thus partitions X into $|B|$ partitions).

$$\begin{tikzcd} X \arrow[r, "f"'] \arrow[rr, bend left, "\chi_{f^{-1}(b)}"] & B \arrow[r, "\chi_{\{b\}}"'] & \mathbf{2} \end{tikzcd} \qquad \begin{tikzcd} \mathrm{Hom}(X, B) \arrow[r, "\alpha_B"] \arrow[d, "(\chi_{\{b\}})_*"'] & B \arrow[d, "\chi_{\{b\}}"] \\ \mathrm{Hom}(X, \mathbf{2}) \arrow[r, "\alpha_{\mathbf{2}}"'] & \mathbf{2} \end{tikzcd}$$

It holds for any $b \in B$ and $x \in X$ that

$$\chi_{\{b\}}(f(x)) = \begin{cases} 1, & f(x) = b \\ 0, & \text{otherwise.} \end{cases}$$

where χ is defined as in equation 8.8.1 of lemma 8.8.0.4.

Since the condition $f(x) = b$ is the same as $b \in f^{-1}(x)$, then

$$\chi_{\{b\}} \circ f = \chi_{f^{-1}(b)}. \tag{8.8.2}$$

By naturality of α, it holds that

$$\alpha_{\mathbf{2}} \circ (\chi_{\{b\}})_* = \chi_{\{b\}} \circ \alpha_B$$

where $(\chi_{\{b\}})_* = \mathrm{Hom}(X, \chi_{\{b\}})$ is again left-composition by $\chi_{\{b\}}$ and so

$$\alpha_{\mathbf{2}}(\chi_{\{b\}} \circ f) = (\alpha_{\mathbf{2}} \circ (\chi_{\{b\}})_*)(f) = (\chi_{\{b\}} \circ \alpha_B)(f). \tag{8.8.3}$$

Define

$$\mathcal{U}_\alpha = \{Y \subseteq X : \alpha_{\mathbf{2}}(\chi_Y) = 1\}$$

where 1 is an element of $\mathbf{2}$.

Then, by definition of $\mathcal{U}_\alpha$, it holds that

$$f^{-1}(b) \in \mathcal{U}_\alpha$$

iff

$$\alpha_{\mathbf{2}}(\chi_{f^{-1}(b)}) = 1$$

which, by equation 8.8.2 happens iff

$$\alpha_{\mathbf{2}}(\chi_{\{b\}} \circ f) = 1$$

which by equation 8.8.3 is true iff

$$\chi_{\{b\}}(\alpha_B(f)) = 1$$

which by definition of χ (equation 8.8.1 of lemma 8.8.0.4) holds iff

$$\alpha_B(f) = b.$$

In summary

$$f^{-1}(b) \in \mathcal{U}_\alpha \qquad \text{iff} \qquad \alpha_B(f) = b$$

which means that $\alpha_B(f)$ is the unique element of B with the property that

$$f^{-1}(\alpha_B(f)) \in \mathcal{U}_\alpha.$$

Hence, if $\mathcal{U}_\alpha$ is an ultrafilter, then

$$\int_X^B f \, \mathrm{d}\mathcal{U}_\alpha = \alpha_B(f).$$

But letting $B = \mathbf{3} = \{0, 1, 2\}$, and $f : X \to \mathbf{3}$ be any map from X to $\mathbf{3}$, there is partition

$$X = X_0 \cup X_1 \cup X_2 = f^{-1}(0) \cup f^{-1}(1) \cup f^{-1}(2)$$

and since $\alpha_{\mathbf{3}}(f)$ is unique with

$$X_{\alpha_{\mathbf{3}}(f)} = f^{-1}(\alpha_{\mathbf{3}}(f)) \in \mathcal{U}_\alpha$$

then $\mathcal{U}_\alpha$ satisfies the 3-partition condition of definition 3.4.0.1 in §3.4 and is hence an ultrafilter by lemma 3.4.0.2 in §3.4. □

Corollary 8.8.0.6. The natural transformation $\int\_\ \mathrm{d} : U \Rightarrow T^G$ is a natural isomorphism.

Proof. At each set $X \in \mathbf{Set}$, the map

$$\int_X \mathrm{d}(-) : U(X) \to T^G(X)$$

is injective and surjective by lemmas 8.8.0.4 and 8.8.0.5.

Hence it is a bijection of sets and thus an isomorphism. So the natural transformation $\int\_\ \mathrm{d}$ is a natural isomorphism. □

Theorem 8.8.0.7. The ultrafilter monad and the codensity monad of $G : \mathbf{FinSet} \hookrightarrow \mathbf{Set}$ are isomorphic, with isomorphism given by the natural isomorphism $\int_X \mathrm{d}(-)$.

Proof. By corollary 3.5.2.6, since T^G is isomorphic to U, they must be isomorphic as monads. □

Remark. The proofs of this section can be adapted for any subcategory $B \subsetneq \mathbf{FinSet}$ that contains at least one set with at least three elements. Hence the codensity monad for the inclusion of any such category B into $\mathbf{Set}$ would also be the ultrafilter monad. For the details, see [Lei13].

8.9 Ulraproducts as Codensity Monads - Introduction

The next half of this chapter concerns a theorem which proves that ultra**products** also arise from a certain codensity monad.

The theorem and a proof originally appear in [Lei13, §8], the author of which credits the proof to an anonymous referee.

The intent of this chapter is to explain some of the necessary prerequisites for understanding the theorem in the above paper, and to provide some intuition as to why the theorem may be true. However, the intuition given is merely an argument for the **plausibility** of the theorem (as opposed to a full proof), and follows a very different path to the proof given in the paper. This plausibility argument was developed for this dissertation.

The theorem is as follows.

Theorem 8.9.0.1. Let $\mathbb{C}$ be a category with all small products and filtered colimits.

The codensity monad for the inclusion functor $\mathbf{FinFam}(\mathbb{C}) \hookrightarrow \mathbf{Fam}(\mathbb{C})$ is isomorphic to the **ultraproduct** monad for $\mathbf{Fam}(\mathbb{C})$.

(See §8.11 for the definition of the ultraproduct monad).

Corollary 8.9.0.2. The codensity monad for $\mathbf{FinFam}(\mathbf{Set}) \hookrightarrow \mathbf{Fam}(\mathbf{Set})$ is isomorphic to the ultraproduct monad for $\mathbf{Fam}(\mathbf{Set})$.

Note, in fact, that varieties of universal algebra with a finite signature (such as rings and groups) satisfy the requirements of the hypothesis. In these categories, the underlying sets of limits and directed colimits are the limits and directed colimits of the underlying sets respectively. Hence, the underlying set of an ultraproduct of a family of such structures is the same as the ultraproduct of the underlying sets of the structures in the family.

In fact, it also works for the category of relational structures (§1.3.5), which gives the general means for obtaining ultraproducts of concrete structures with finite signature (even in 'badly behaved' categories such as fields).

8.10 Category of Families of Structures

Definition 8.10.0.1. Fix some category $\mathbb{C}$ with small products and filtered colimits.

The **category $\mathbf{Fam}(\mathbb{C})$ of $\mathbb{C}$-families** is defined as follows.

- An object
$$\langle X, \langle S_x \rangle_{x \in X} \rangle$$
is a pair consisting of
  - an indexing set X.
  - a family $\langle S_x \rangle_{x \in X}$ of objects $S_x \in \mathbb{C}$ indexed by X.

  This will often be denoted just as $\langle S_x \rangle_{x \in X}$.

- A morphism

$$\langle f, \langle \phi_x \rangle_{x \in X} \rangle : \langle X, \langle S_x \rangle_{x \in X} \rangle \to \langle Y, \langle R_y \rangle_{y \in Y} \rangle$$

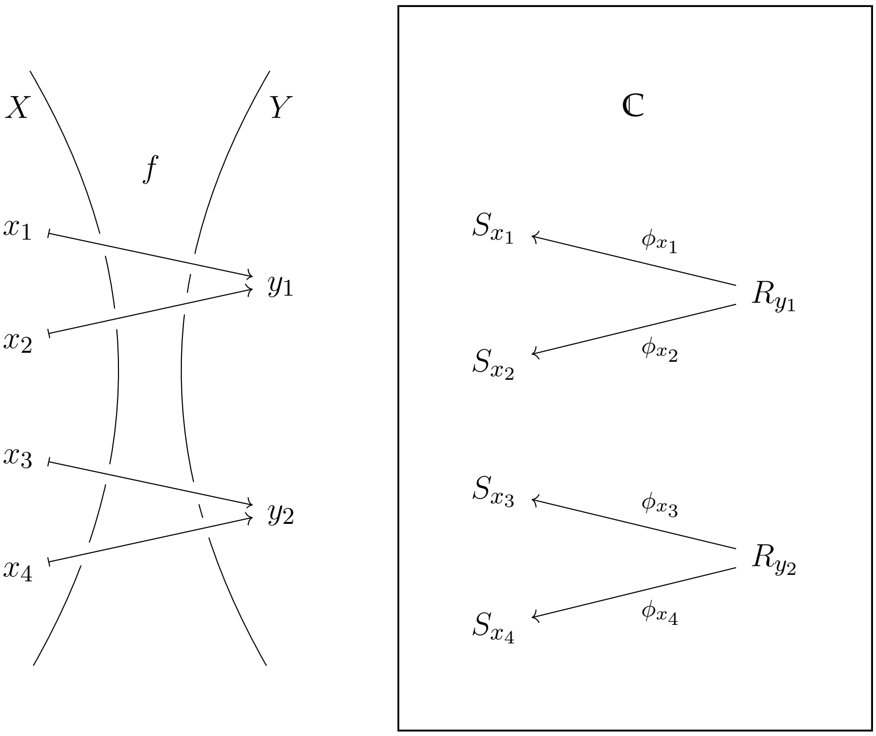
  is a pair consisting of

  - a map $f : X \to Y$.
  - a family of morphisms $\left\langle \phi_x : R_{f(x)} \to S_x \right\rangle_{x \in X}$ in $\mathbb{C}$.

◁

Remark. The notation for the family of morphisms $\langle \phi_x \rangle_{x \in X}$ is the same as that for an induced morphism into a product. However, the usage should be clear from context. ◁

Remark. Note that the morphisms ϕ_x go 'backward' $R_{f(x)} \to S_x$ instead of vice-versa. This follows the approach of [Lei13], but may differ from some texts. ◁



Definition 8.10.0.2. Given $\mathbf{Fam}(\mathbb{C})$, the **subcategory** $\mathbf{FinFam}(\mathbb{C})$ **of finite** $\mathbb{C}$**-families** is the full subcategory where objects are those families $\langle S_x \rangle_{x \in X}$ where the indexing set X is finite. ◁

8.11 Ultraproduct Monad

The functor for the ultraproduct monad is a functor V which maps objects by

$$V : \mathbf{Fam}(\mathbb{C}) \to \mathbf{Fam}(\mathbb{C}) : \langle S_x \rangle_{x \in X} \mapsto \left\langle \prod_{\mathcal{U}} S_x \right\rangle_{\mathcal{U} \in U(X)}$$

and maps morphisms by

$$\mathbf{Fam}(\mathbb{C}) \xrightarrow{\quad V \quad} \mathbf{Fam}(\mathbb{C})$$

$$\begin{array}{ccc} \langle S_x \rangle_{x \in X} & & \left\langle \prod_{\mathcal{U}} S_x \right\rangle_{\mathcal{U} \in U(X)} \\ \Big\downarrow {\scriptstyle \langle f \,,\, \langle \phi_x \rangle_{x \in X} \rangle} & & \Big\downarrow {\scriptstyle \langle f_* \,,\, \langle \psi_{\mathcal{U}} \rangle_{\mathcal{U} \in U(X)} \rangle} \\ \langle R_y \rangle_{y \in Y} & & \left\langle \prod_{\mathcal{V}} R_y \right\rangle_{\mathcal{V} \in U(Y)} \end{array}$$

where the image of the morphism is a pair

$$\left\langle f_* : U(X) \to U(Y) \quad , \quad \left\langle \psi_{\mathcal{U}} : \prod_{f_*(\mathcal{U})} R_y \to \prod_{\mathcal{U}} S_x \right\rangle_{\mathcal{U} \in U(X)} \right\rangle$$

The first component of the pair is defined by $f_* = U(f)$ as defined in §3.5.1. As a reminder, it is

$$U(f) : U(X) \to U(Y) : \mathcal{U} \mapsto f_*\mathcal{U} = \{Z \subseteq Y : f^{-1}(Z) \in \mathcal{U}\}.$$

The second component is more complex.

Recall that for ultrafilter $\mathcal{U} \subseteq \mathcal{P}(X)$, the ultraproduct is given by

$$\prod_{\mathcal{U}} S_x = \varinjlim_{H \in \mathcal{U}} \prod_{x \in H} S_x,$$

and for ultrafilter $f_*(\mathcal{U}) \subseteq \mathcal{P}(Y)$, the ultraproduct is given by

$$\prod_{f_*(\mathcal{U})} R_y = \varinjlim_{K \in f_*(\mathcal{U})} \prod_{y \in K} R_y.$$

Then each $\psi_{\mathcal{U}}$ is a morphism

$$\varinjlim_{K\in f_*(\mathcal{U})} \prod_{y\in K} R_y \xrightarrow{\psi_{\mathcal{U}}} \varinjlim_{H\in\mathcal{U}} \prod_{x\in H} S_x .$$

It is defined in terms of the f and ϕ_x morphisms.

A morphism out of a colimit is induced by a family of morphisms out of its components. In the commutative diagram below, the dashed arrows are induced by morphisms below them, and the dotted arrows are compositions of morphisms below them.

$$\begin{tikzcd}
\varinjlim\limits_{K\in f_*(\mathcal{U})} \prod\limits_{y\in K} R_y \arrow[rr, dashed, "\psi_{\mathcal{U}}"] & & \varinjlim\limits_{H\in\mathcal{U}} \prod\limits_{x\in H} S_x \\
\prod\limits_{y\in K} R_y \arrow[u, "\gamma"] \arrow[r, dashed] \arrow[urr, dotted] \arrow[dr, dotted] \arrow[d, "\pi"'] & \prod\limits_{y\in K} \prod\limits_{x\in f^{-1}(y)} S_x \arrow[r, tail, two heads, "\text{iso}"] \arrow[d, "\pi"'] & \prod\limits_{x\in f^{-1}(K)} S_x \arrow[u, "\gamma"'] \\
R_{y=f(x)} \arrow[r, dashed] \arrow[dr, "\phi_x"'] & \prod\limits_{x\in f^{-1}(y)} S_x \arrow[d, "\pi"'] & \\
& S_x &
\end{tikzcd}$$

The same diagram is displayed another way below. The greyed-out objects and arrows are other examples of (co)projections in the (co)limit.

$$\begin{tikzcd}
& \varinjlim\limits_{K\in f_*(\mathcal{U})} \prod\limits_{y\in K} R_y \arrow[rr, dashed, "\psi_{\mathcal{U}}"] & & \varinjlim\limits_{H\in\mathcal{U}} \prod\limits_{x\in H} S_x & \\
\prod\limits_{y\in K'} R_y \arrow[ur] & \prod\limits_{y\in K} R_y \arrow[u] \arrow[rr, dashed, "\langle \phi_x\ \pi_{f(x)} \rangle_{x\in f^{-1}(K)}"] \arrow[ddl] \arrow[dd] & & \prod\limits_{x\in f^{-1}(K)} S_x \arrow[u] \arrow[dl] \arrow[ddl] \arrow[dddl] & \prod\limits_{x\in H} S_x \arrow[ul] \arrow[dd] \\
& & S_{x_1} & & \\
R_{y'} & R_y \arrow[ur, "\phi_{x_1}"] \arrow[r, "\phi_{x_2}"] \arrow[dr, "\phi_{x_3}"'] & S_{x_2} & & S_{x'} \\
& & S_{x_3} & &
\end{tikzcd}$$

8.12 Category Theory on Families of Structures

Now, given a specific set X, one can view it as a discrete category. Then a family $\langle S_x \rangle_{x\in X}$ of $\mathbb{C}$ structures can be seen as a functor $S : X \to \mathbb{C}$, where $S_x = S(x)$ for $x \in X$.

This functor S is an element of the functor category $\mathbb{C}^X$, whose morphisms are natural transformations $\phi : S \Rightarrow R$. Such a natural transformation from a family $\langle S_x \rangle_{x\in X}$ to $\langle R_x \rangle_{x\in X}$ is a family of morphisms $\langle \phi_x : S_x \to R_x \rangle_{x\in X}$.

Remark. To maintain consistency with the rest of the chapter, these should actually be reversed. The family $\langle \phi_x : R_x \to S_x \rangle_{x\in X}$ and hence the category $\left(\mathbb{C}^X\right)^{\mathrm{op}}$ should be used. Hence a morphism $\phi : S \to R$ corresponds to a natural transformation $\phi : R \Rightarrow S$. ◁

Given a morphism $f : X \to Y$ (which can be seen as a functor between discrete categories X and Y), and a family $R : Y \to \mathbb{C}$, one can obtain a family $Rf : X \to \mathbb{C}$. In more detail, the family $\langle R_y \rangle_{y\in Y}$ is mapped to $\langle R_{f(x)} \rangle_{x\in X}$.

Additionally, given a family of morphisms $\langle \phi_y : R'_y \to R_y \rangle_{y\in Y}$ from a family $\langle R_y \rangle_{y\in Y}$ to family $\langle R'_y \rangle_{y\in Y}$ (or, equivalently, a natural transformation $\phi : R' \Rightarrow R$), one obtains $\phi f = \langle \phi_{f(x)} : R'_f(x) \to R_{f(x)} \rangle_{x\in X}$ from Rf to $R'f$.

One then obtains a functor

$$f^* : \quad \left(\mathbb{C}^Y\right)^{\mathrm{op}} \to \left(\mathbb{C}^X\right)^{\mathrm{op}} : \quad R \mapsto Rf, \quad \phi \mapsto \phi f$$

mapping objects and morphisms as above.

It is also possible to define a functor

$$f_* : \left(\mathbb{C}^X\right)^{\mathrm{op}} \to \left(\mathbb{C}^Y\right)^{\mathrm{op}} : \langle S_x \rangle_{x\in X} \mapsto \left\langle \prod_{x\in f^{-1}(y)} S_x \right\rangle_{y\in Y}$$

which maps a morphism

$$\langle \phi_x : S'_x \to S_x \rangle_{x\in X} : S \to S'$$

to the family of induced morphisms between products

$$\left\langle \langle \phi_x \pi_{f(x)} \rangle_{x\in f^{-1}(y)} : \prod_{x\in f^{-1}(y)} S'_x \to \prod_{x\in f^{-1}(y)} S_x \right\rangle_{y\in Y}$$

which is itself a morphism between

$$\left\langle \prod_{x\in f^{-1}(y)} S_x \right\rangle_{y\in Y} \to \left\langle \prod_{x\in f^{-1}(y)} S'_x \right\rangle_{y\in Y}.$$

As it turns out, f_* is left adjoint to f^*.

$$
\begin{tikzcd}
\left(\mathbb{C}^X\right)^{\mathrm{op}} \arrow[rr, bend left, "f_*"] \arrow[rr, phantom, "\perp"] & & \left(\mathbb{C}^Y\right)^{\mathrm{op}} \arrow[ll, bend left, "f^*"]
\end{tikzcd}
\tag{8.12.1}
$$

It is possible to see this by examining the natural bijection of the adjunction.

$$
\begin{tikzcd}
\langle S_x \rangle_{x\in X} \arrow[r, mapsto, "f_*"] \arrow[d, "\langle R_{f(x)} \to S_x \rangle_{x\in X}"] & \left\langle \prod\limits_{x\in f^{-1}(y)} S_x \right\rangle_{y\in Y} \\
\langle R_{f(x)} \rangle_{x\in X} & \langle R_y \rangle_{y\in Y} \arrow[l, mapsto, "f^*"'] \arrow[u, leftarrow, "\left\langle R_y \to \prod\limits_{x\in f^{-1}(y)} S_x \right\rangle_{y\in Y}"']
\end{tikzcd}
$$

where the forwards direction of the bijection takes a family of morphisms

$$\langle R_{f(x)} \to S_x \rangle_{x\in X}$$

to a family

$$\left\langle R_y \to \prod_{x\in f^{-1}(y)} S_x \right\rangle_{y\in Y}$$

and the backwards direction takes it back.

The properties of bijectivity and naturality then follow directly from the universal property of products

This is illustrated, by example, in the following diagram, where x_1, x_2 and x_3 are all x such that $y = f(x)$. Then there is an induced morphism (illustrated by the dashed line), and there is a bijection between such induced morphisms and cones to the S_x from R_y.

$$
\begin{tikzcd}
 & & S_{x_1} \\
R_{y=f(x)} \arrow[r, dashed] \arrow[rru, bend left] \arrow[rr, bend left] \arrow[rrd, bend right] & \prod\limits_{x\in f^{-1}(y)} S_x \arrow[ru] \arrow[r] \arrow[rd] & S_{x_2} \\
 & & S_{x_3}
\end{tikzcd}
$$

There is also the projection functor

$$\mathrm{pr} : \mathbb{C}^{\mathbf{Set}} \to \mathbf{Set}$$

which takes a family $\langle S_x \rangle_{x \in X}$ to its indexing set X. There is a similar functor

$$\mathrm{pr} : \mathbb{C}^{\mathbf{FinSet}} \to \mathbf{FinSet}.$$

8.13 Intuition for Why End Leads to Ultraproducts

In this section, some intuition is given for why the specific end arising from the codensity monad is, in fact, the family of ultraproducts indexed by ultrafilters.

In particular, the details of the morphisms out of the end are given and a brief description of how commutativity and universality for the end follow from commutativity and universality for the ultraproduct. It is emphasized that this is not a proof, but merely an argument for why the ultraproduct family being the end is **plausible**. A full proof may be found in [Lei13], but follows a very different approach.

It will be convenient to refer to families in pair notation $\langle X, S \rangle = \langle S_x \rangle_{x \in X}$, where S is a functor $S : X \to \mathbb{C}$, considered in $\left(\mathbb{C}^X\right)^{\mathrm{op}}$.

Now, given a family $\langle X, S \rangle$, the family of ultraproducts yields an end. Specifically, it is the universal family $\langle U(X)\ ,\ \prod_{-} S_x \rangle$ and family of morphisms $\chi_{\langle B,R \rangle}$ making the following diagram commute for each morphism $\langle g, \psi \rangle$.

The details of how this diagram relates to the codensity monad are given in §8.13.3.

$$\begin{tikzcd}
 & \prod\limits_{\mathrm{Hom}(\langle X,S\rangle,\langle B,R\rangle)} \langle B, R\rangle \arrow[d, "\left\langle \langle g,\psi\rangle \pi^{\langle B,R\rangle}_{\langle f,\phi\rangle} \right\rangle_{\langle f,\phi\rangle \in \mathrm{Hom}(\langle X,S\rangle,\langle B,R\rangle)}"] \\
\left\langle U(X)\ ,\ \prod\limits_{-} S_x \right\rangle \arrow[ur, "\chi_{\langle B,R\rangle}"] \arrow[dr, "\chi_{\langle B',R'\rangle}"'] & \prod\limits_{\mathrm{Hom}(\langle X,S\rangle,\langle B,R\rangle)} \langle B', R'\rangle \\
 & \prod\limits_{\mathrm{Hom}(\langle X,S\rangle,\langle B',R'\rangle)} \langle B', R'\rangle \arrow[u, "\left\langle \pi^{\langle B',R'\rangle}_{\langle gf,\psi\phi\rangle} \right\rangle_{\langle f,\phi\rangle \in \mathrm{Hom}(\langle X,S\rangle,\langle B,R\rangle)}"']
\end{tikzcd}$$

The map $\chi_{\langle B,R \rangle}$ is induced by components

$$\chi^{\langle f,\phi\rangle}_{\langle B,R\rangle} : \left\langle U(X)\ ,\ \prod_{-} S_x \right\rangle \to \langle B, R \rangle$$

each of which is a pair

$$\left\langle \int_X^B f \ \mathrm{d}- \ , \ \lambda^\phi \right\rangle : \left\langle U(X) \ , \ \prod_{-} S_x \right\rangle \to \langle B, R \rangle$$

where the map

$$\int_X^B f \ \mathrm{d}- : U(X) \to B$$

works exactly as in §8.8 and the λ^ϕ is a family of maps

$$\left\langle \lambda^\phi_{\mathcal{U}} : R_{\int_X^B f \ \mathrm{d}\mathcal{U}} \to \prod_{\mathcal{U}} S_x \right\rangle_{\mathcal{U} \in U(X)}$$

defined such that the following diagram is commutative, where $H = f^{-1}(\int_X^B f \ \mathrm{d}\mathcal{U})$, (and $H \in \mathcal{U}$ by definition)

$$\begin{tikzcd}
 & \prod_{\mathcal{U}} S_x \\
R_{\int_X^B f \ \mathrm{d}\mathcal{U}} \arrow[ur, "\lambda^\phi_{\mathcal{U}} = \text{ш}_H \ \langle \phi_x \rangle_{x \in H}"] \arrow[r, "\langle \phi_x \rangle_{x \in H}"] \arrow[dr, "\phi_x"'] & \prod_{x \in H = f^{-1}(\int_X^B f \ \mathrm{d}\mathcal{U})} S_x \arrow[u, "\text{ш}_H"'] \arrow[d, "\pi_x"] \\
 & S_x
\end{tikzcd}$$

Why should this be the case?

Consider an arbitrary $\langle f, \phi \rangle$.

The $\int_X^B f \ \mathrm{d}\mathcal{U} \in B$ picks out a specific element b in the finite set B, (namely the one whose inverse image under f is $\mathcal{U}$-large). Since B indexes R, this corresponds to a specific object R_b in the family.

A morphism

$$\lambda^\phi_{\mathcal{U}} : R_b \to \prod_{\mathcal{U}} S_x$$

is completely determined by a family of morphisms $\phi_x : R_b \to S_x$.

The ultraproduct $\prod_{\mathcal{U}} S_x$ is an object satisfying this property for any family $\langle B, R \rangle$.

8.13.1 Commutativity

Furthermore, the ultraproduct family satisfies a commutativity condition in the following sense: For any family $\langle B', R' \rangle$ and morphism

$$\langle g, \psi \rangle : \langle B, R \rangle \to \langle B', R' \rangle$$

for each $\mathcal{U} \in U(X)$, and where again $b = \int_X^B f \; \mathrm{d}\mathcal{U}$, it holds that

$$\left\langle g \left(\int_X^B f \; \mathrm{d}\mathcal{U} \right) \; , \; \lambda_{\mathcal{U}}^{\phi} \, \psi_b \right\rangle \; = \; \left\langle \left(\int_X^{B'} gf \; \mathrm{d}\mathcal{U} \right) \; , \; \lambda_{\mathcal{U}}^{\phi\psi} \right\rangle$$

$$\begin{tikzcd}
 & & \prod_{\mathcal{U}} S_x \\
R'_{g(b)=gf(x)} \arrow[r, "\psi_b"] \arrow[rru, bend left, "\lambda_{\mathcal{U}}^{\phi\psi}"] \arrow[rrd, bend right, "\phi_x \psi_{f(x)} = (\phi\psi)_x"'] & R_b \arrow[r, "\langle \phi_x \rangle_{x \in H}"] \arrow[ru, "\lambda_{\mathcal{U}}^{\phi}"] \arrow[rd, "\phi_x"] & \prod_{x \in H} S_x \arrow[u, "\rotatebox[origin=c]{180}{π}_H"'] \arrow[d, "\pi_x"'] \\
 & & S_x
\end{tikzcd}$$

where $\lambda_{\mathcal{U}}^{\psi\phi}$ is a component of $\lambda^{\psi\phi}$ which is itself a component of $\chi_{\langle B', R' \rangle}^{\langle gf, \psi\phi \rangle}$ which is, again, itself a component of $\chi_{\langle B', R' \rangle}$.

What this represents firstly, is the fact as proven in [Lei13] that if f and g are composed, then the unique element $b' \in B'$ whose inverse image under gf is in the ultrafilter, is just the image under g of the unique element of B whose inverse image under f is in the ultrafilter.

Furthermore, a morphism from the unique $R'_{b'}$ into the ultraproduct given by composing with ψ_b the morphism $\lambda_{\mathcal{U}}^{\phi}$ from R_b into the ultraproduct is the morphism into the ultraproduct that is induced by composing each of the relevant ϕ_x morphisms with ψ_b.

$$\begin{tikzcd}
S_{x_1} & & \\
& R_b \arrow[ul, "\phi_{x_1}"'] \arrow[dl, "\phi_{x_2}"] & \\
S_{x_2} & & \\
& & R'_{g(b)} \arrow[uul, "\psi_b"'] \arrow[ddl, "\psi_{b_2}"] \\
S_{x_3} & & \\
& R_{b_2} \arrow[ul, "\phi_{x_3}"'] \arrow[dl, "\phi_{x_4}"] & \\
S_{x_4} & &
\end{tikzcd}$$

8.13.2 Universality

Let $\langle A, Q\rangle$ be a family, and ξ be its family of morphisms satisfying the commutativity of the square for the end.

For the family of ultraproducts to be an end, it is necessary that there is a unique morphism $\langle h, \zeta\rangle$ from this family into the ultraproduct family, such that $(\chi_{\langle B,R\rangle}\ \langle h, \zeta\rangle = \xi_{\langle B,R\rangle}$.

$$\begin{tikzcd}
& & \prod\limits_{\mathrm{Hom}(\langle X,S\rangle,\langle B,R\rangle)} \langle B, R\rangle \arrow[d] \\
\langle A, Q\rangle \arrow[r, dashed, "\langle h,\zeta\rangle"] \arrow[urr, bend left, "\xi_{\langle B,R\rangle}"] \arrow[drr, bend right, "\xi_{\langle B',R'\rangle}"'] & \left\langle U(X)\ ,\ \prod\limits_{-} S_x \right\rangle \arrow[ur, "\chi_{\langle B,R\rangle}"] \arrow[dr, "\chi_{\langle B',R'\rangle}"'] & \prod\limits_{\mathrm{Hom}(\langle X,S\rangle,\langle B,R\rangle)} \langle B', R'\rangle \\
& & \prod\limits_{\mathrm{Hom}(\langle X,S\rangle,\langle B',R'\rangle)} \langle B', R'\rangle \arrow[u]
\end{tikzcd}$$

Let the components of ξ be given as

$$\xi^{\langle f,\phi\rangle}_{\langle B,R\rangle} = \langle \alpha^f,\ \delta^\phi\rangle : \langle A, Q\rangle \to \langle B, R\rangle$$

where $\alpha^f : A \to B$ is a map and δ^ϕ is a family of morphisms

$$\langle \delta^\phi_a : R_{\alpha^f(a)} \to Q_a\rangle_{a\in A}.$$

Then, because it was shown in [Lei13] that the ultra**filter** monad is given as the relevant end, then it immediately follows there is a unique $h : A \to U(X)$

such that if $h(a) = \mathcal{U}$ it holds that

$$\alpha^f(a) = \int_X^B f \; \mathrm{d}\mathcal{U}.$$

Recall that a morphism

$$\langle h, \zeta \rangle : \langle A, Q \rangle \to \left\langle U(X) \, , \, \prod_{-} S_x \right\rangle$$

has a family of morphisms for each $a \in A$

$$\zeta_a : \prod_{\mathcal{U}=h(a)} S_x \to Q_a$$

going **out** of each ultraproduct.

Hence, such a ζ_a may be obtained by showing that each component Q_a of Q has a family of morphisms δ_a^ϕ which form a cone over the products from which the ultraproduct is composed. This should be shown using the commutativity of the ξ morphisms in the end diagram.

Once that is done, it remains only to show that the ζ_a so obtained form the components of the ζ morphism, by showing that the correct commutativity is satisfied, and that this ζ is the unique morphism which works.

The following diagram is intended as an illustration of the above, for two sets $K, H \in \mathcal{U}$ with $K \subseteq H$. All of the relevant morphisms commute.

$$\begin{tikzcd}
 & & Q_a \\
 & & \prod_{\mathcal{U}} S_x \arrow[u, dashed, "\zeta_a"] \\
R'_{b'} = \prod_{x \in H} S_x \arrow[rruu, bend left, "\delta_a^{\phi\psi}"] \arrow[rru, bend left, "\lambda_{\mathcal{U}}^{\phi\psi}"] \arrow[rr, "\pi_K^H"] \arrow[rrd, "\pi_x"] & & \prod_{x \in K} S_x = R_{b'} \arrow[u, "\amalg_K = \lambda_{\mathcal{U}}^{\phi}"] \arrow[uu, bend right, "\delta_a^{\phi}"'] \arrow[d, "\pi_x"] \\
 & & S_x
\end{tikzcd}$$

A proof is not given here.

8.13.3 Derivation of the End in More Detail

Let

$$G' : \mathbf{FinFam}(\mathbb{C}) \hookrightarrow \mathbf{Fam}(\mathbb{C})$$

be the functor which includes finite families into families.

As shown in [Lei13], and also eariler in this dissertation, there is an adjunction.

$$\begin{tikzcd} \mathbf{Fam}(\mathbb{C}) \arrow[rr, bend left, "\mathrm{Hom}(-{,}G')"] \arrow[rr, phantom, "\perp"] & & \left(\mathbf{Set}^{\mathbf{FinFam}(\mathbb{C})}\right)^{\mathrm{op}} \arrow[ll, bend left, "\mathrm{Nat}(-{,}G')"] \end{tikzcd}$$

Further, it was shown that for a functor $F : \mathbf{FinFam}(\mathbb{C}) \to \mathbf{Set}$, the Nat functor can be described as an end by

$$\mathrm{Nat}(F, G') = \int_{\langle B,R\rangle \in \mathbf{FinFam}(\mathbb{C})} \left(\prod_{F(\langle B,R\rangle)} G'(\langle B, R\rangle) \right)$$

where

$$\prod_{F(-)} G'(*) : \mathbf{FinFam}(\mathbb{C})^{\mathrm{op}} \times \mathbf{FinFam}(\mathbb{C}) \to \mathbf{Fam}(\mathbb{C}).$$

The first component (the dash) comes from $\mathbf{FinFam}(\mathbb{C})^{\mathrm{op}}$ and the second (the asterisk) from $\mathbf{FinFam}(\mathbb{C})$.

This functor maps a pair $\langle\langle C, S\rangle, \langle B, R\rangle\rangle$ to

$$\prod_{F(\langle C,S\rangle)} G'(\langle B, R\rangle)$$

and a pair of morphisms

$$\langle\langle g', \psi'\rangle, \langle g, \psi\rangle\rangle : \langle\langle C', S'\rangle, \langle B, R\rangle\rangle \to \langle\langle C, S\rangle, \langle B', R'\rangle\rangle$$

where

$$\langle g', \psi'\rangle : \langle C, S\rangle \to \langle C', S'\rangle$$

and

$$\langle g, \psi\rangle : \langle B, R\rangle \to \langle B', R'\rangle$$

to

$$\prod_{F(\langle g',\psi'\rangle)} G'(\langle g, \psi\rangle) : \prod_{F(\langle C',S'\rangle)} G'(\langle B, R\rangle) \to \prod_{F(\langle C,S\rangle)} G'(\langle B', R'\rangle)$$

where

$$\prod_{F(\langle g',\psi'\rangle)} G'(\langle g, \psi\rangle) = \left\langle G'(\langle g, \psi\rangle)\pi^{G'(\langle B,R\rangle)}_{F(\langle g',\psi'\rangle)(x)} \right\rangle_{x \in F(\langle C,S\rangle)}$$

For notational simplicity, the inclusion can be ignored.

$$\mathrm{Nat}(F, G') = \int_{\langle B,R\rangle \in \mathbf{FinFam}(\mathbb{C})} \left(\prod_{F(\langle B,R\rangle)} \langle B, R\rangle \right)$$

This has a family of morphisms

$$\left\langle \chi_{\langle B,R\rangle} : \mathrm{Nat}(F, G') \to \prod_{F(\langle B,R\rangle)} \langle B, R\rangle \right\rangle_{\langle B,R\rangle \in \mathbf{FinFam}(\mathbb{C})}$$

where for every $\langle B, R\rangle, \langle B', R'\rangle \in \mathbf{FinFam}(\mathbb{C})$ and $\langle g, \psi\rangle : \langle B, R\rangle \to \langle B', R'\rangle$ the following diagram commutes:

$$\begin{tikzcd}
& \prod\limits_{F(\langle B,R\rangle)} \langle B, R\rangle \arrow[dr, "\left\langle \langle g,\psi\rangle \pi_x^{\langle B,R\rangle} \right\rangle_{x \in F(\langle B,R\rangle)}"] & \\
\mathrm{Nat}(F, G') \arrow[ur, "\chi_{\langle B,R\rangle}"] \arrow[dr, "\chi_{\langle B',R'\rangle}"'] & & \prod\limits_{F(\langle B,R\rangle)} \langle B', R'\rangle \\
& \prod\limits_{F(\langle B',R'\rangle)} \langle B', R'\rangle \arrow[ur, "\left\langle \pi^{\langle B',R'\rangle}_{F(\langle g,\psi\rangle)(x)} \right\rangle_{x \in F(\langle B,R\rangle)}"'] &
\end{tikzcd}$$

In particular, the functor $T^{G'} : \mathbf{Fam}(\mathbb{C}) \to \mathbf{Fam}(\mathbb{C})$ of the codensity monad is defined on each family $\langle X, S\rangle$ by

$$T^{G'}(\langle X, S\rangle) = \int_{\langle B,R\rangle \in \mathbf{FinFam}(\mathbb{C})} \left(\prod_{\mathrm{Hom}(\langle X,S\rangle, \langle B,R\rangle)} \langle B, R\rangle \right)$$

This has a family of morphisms

$$\left\langle \chi_{\langle B,R\rangle} : \mathrm{Nat}(\mathrm{Hom}(\langle X, S\rangle, -), G') \to \prod_{\mathrm{Hom}(\langle X,S\rangle, \langle B,R\rangle)} \langle B, R\rangle \right\rangle_{\langle B,R\rangle \in \mathbf{FinFam}(\mathbb{C})}$$

where for every $\langle B, R\rangle, \langle B', R'\rangle \in \mathbf{FinFam}(\mathbb{C})$ and $\langle g, \psi\rangle : \langle B, R\rangle \to \langle B', R'\rangle$ the following diagram commutes:

$$\begin{tikzcd}[column sep=large] & \prod_{\mathrm{Hom}(\langle X,S\rangle,\langle B,R\rangle)} \langle B,R\rangle \arrow[d, "\left\langle \langle g,\psi\rangle \pi^{\langle B,R\rangle}_{\langle f,\phi\rangle} \right\rangle_{\langle f,\phi\rangle \in \mathrm{Hom}(\langle X,S\rangle,\langle B,R\rangle)}"] \\ \mathrm{Nat}(\mathrm{Hom}(\langle X,S\rangle,-),G') \arrow[ur, "\chi_{\langle B,R\rangle}"] \arrow[dr, "\chi_{\langle B',R'\rangle}"'] & \prod_{\mathrm{Hom}(\langle X,S\rangle,\langle B,R\rangle)} \langle B',R'\rangle \\ & \prod_{\mathrm{Hom}(\langle X,S\rangle,\langle B',R'\rangle)} \langle B',R'\rangle \arrow[u, "\left\langle \pi^{\langle B',R'\rangle}_{\langle gf,\psi\phi\rangle} \right\rangle_{\langle f,\phi\rangle \in \mathrm{Hom}(\langle X,S\rangle,\langle B,R\rangle)}"'] \end{tikzcd}$$

This is given by

$$\mathrm{Nat}(\mathrm{Hom}(\langle X,S\rangle,-),G') = \left\langle \prod_{\mathcal{U}} S_x \right\rangle_{\mathcal{U}\in U(X)} = \left\langle U(X) \quad , \quad \prod_{-} S_x \right\rangle$$

Each of the χ maps is

$$\chi_{\langle B,R\rangle} : \left\langle U(X) \quad , \quad \prod_{-} S_x \right\rangle \to \prod_{\mathrm{Hom}(\langle X,S\rangle,\langle B,R\rangle)} \langle B,R\rangle$$

They are the morphisms into the products induced by the morphisms, for each

$$\langle f,\phi\rangle : \langle X,S\rangle \to \langle B,R\rangle$$

given by

$$\left\langle \int_X^B f \ \mathrm{d}- \quad , \quad \lambda^{\phi} \right\rangle : \left\langle U(X) \quad , \quad \prod_{-} S_x \right\rangle \to \langle B,R\rangle$$

The map $\int_X^B f \ \mathrm{d}- : U(X) \to B$ maps $\mathcal{U}$ to $\int_X^B f \ \mathrm{d}\mathcal{U}$ (which is the unique $b \in B$ such that $f^{-1}(b) \in \mathcal{U}$).

The λ^{ϕ} is a family of maps

$$\left\langle \lambda^{\phi}_{\mathcal{U}} : R_{\int_X^B f \ \mathrm{d}\mathcal{U}} \to \prod_{\mathcal{U}} S_x \right\rangle_{\mathcal{U}\in U(X)}$$

Recall that ϕ is a family of maps

$$\left\langle \phi_x : R_{f(x)} \to S_x \right\rangle_{x\in X}$$

There is an induced map to the product

$$\langle \phi_x \rangle_{x\in f^{-1}(b)} : R_b \to \prod_{x\in f^{-1}(b)} S_x$$

If $b = \int_X^B f \;\mathrm{d}\mathcal{U}$, then the set $\{x \in X : x \in f^{-1}(b)\} \in \mathcal{U}$ so there is a coprojection

$$\amalg_{f^{-1}(b)} : \prod_{x \in f^{-1}(b)} S_x \to \prod_{\mathcal{U}} S_x$$

Each morphism $\lambda_{\mathcal{U}}^{\phi}$ is then given as the composition

$$\lambda_{\mathcal{U}}^{\phi} = \amalg_{f^{-1}(b)} \circ \langle \phi_x \rangle_{x \in f^{-1}(b)}.$$

And so

$$\chi_{\langle B,R \rangle} := \left\langle \left\langle \int_X^B f \;\mathrm{d}- \quad , \quad \lambda^{\phi} \right\rangle \right\rangle_{\mathrm{Hom}(\langle X,S \rangle, \langle B,R \rangle)}$$

References

[AM69] M. Atiyah and I. G. Macdonald. **Introduction to Commutative Algebra**. Addison-Wesley Publishing, Reading, Massachusetts, first edition, 1969.

[AN78] H. Andreka and I. Németi. łoś lemma holds in every category. **Studia Scientiarum Mathaticarum Hungarica**, 13(3-4):361–376, 1978.

[AN79] H. Andreka and I. Németi. Injectivity in categories to represent all first order formulas, i. **Demonstratio Mathematica**, 12(3):717–732, 1979.

[AN82] H. Andreka and I. Németi. Direct limits and filtered colimits are strongly equivalent in all categories. **Universal Algebra and Applications**, 9:75–88, 1982.

[Ax68] J. Ax. The elementary theory of finite fields. **Annals of Mathematics**, 88(2):239–271, 1968.

[Bar77a] J. Barwise. **Handbook of Mathematical Logic**. North-Holland Publishing Company, Amsterdam, New York, Oxford, first edition, 1977.

[Bar77b] J. Barwise. An introduction to first-order logic. In J. Barwise, editor, **Handbook of Mathematical Logic**, chapter A1, pages 5–46. North-Holland Publishing Company, Amsterdam, New York, Oxford, first edition, 1977.

[BM77] J.L. Bell and M Machover. **A Course in Mathematical Logic**. North-Holland Publishing Company, Amsterdam, New York, Oxford, first edition, 1977.

[BN87] H. H. Bui and I. Németi. Problems with the category theoretic notions of ultraproducts. **Bulletin of the Section of Logic**, 10(3):122–127, 1987.

[Bör87] R. Börger. Coproducts and ultrafilters. **Journal of Pure and Applied Algebra**, 46:35–47, 1987.

[Bor94] F. Borceux. Basic category theory. In F. Borceux, editor, **Handbook of Categorical Algebra**, volume 1. Cambridge University Press, Cambridge, 1994.

[BS71] J.L. Bell and A.B. Slomson. **Models and Ultraproducts: An Introduction**. North-Holland Publishing Company, Amsterdam, second revised edition, 1971.

[BS81] S. Burris and H.P. Sankappanavar. A course in universal algebra. `https://www.math.uwaterloo.ca/~snburris/htdocs/UALG/univ-algebra.pdf` retrived 2016, original 1981.

[CK73] C. C. Chang and H.J. Keisler. Studies in logic and the foundations of mathematics. In A. Heyting et. al, editor, **Model Theory**, volume 73. North-Holland Publishing Company, Amsterdam, New York, Oxford, 1973.

[CN70] W.W. Comfort and S. Negrepontis. **The Theory of Ultrafilters**. Springer-Verlag, Berlin Heidelberg New York, first edition, 1970.

[DP02] B.A. Davey and H.A. Priestley. **Introduction to Lattices and Order**. Cambridge University Press, Cambridge, second edition, 2002.

[Ekl77] P.C. Eklof. Ultraproducts for algebraists. In J. Barwise, editor, **Handbook of Mathematical Logic**, chapter A3, pages 105–138. North-Holland Publishing Company, Amsterdam, New York, Oxford, first edition, 1977.

[GH70] F. Galvin and A. Horn. Operations preserving all equivalence relations. **Proceedings of the American Mathematical Society**, 24(3):521–523, 1970.

[Her06] H. Herrlich. **Axiom of Choice**. Springer-Verlag, Berlin, Heidelberg, first edition, 2006.

[HMT71] L. Henkin, J.D. Monk, and A. Taski. **Cylindric Algebras, Part I**. North-Holland Publishing Company, Amsterdam, first edition, 1971.

[HS73] H. Herrlich and E.S. Strecker. **Category Theory**. Allyn and Bacon, Inc., Atlantic Avenue, Boston, first edition, 1973.

[Jec77] T.J. Jech. About the axiom of choice. In J. Barwise, editor, **Handbook of Mathematical Logic**, chapter B2, pages 345–370. North-Holland Publishing Company, Amsterdam, New York, Oxford, first edition, 1977.

[JHS05] A. Jiří, H. Herrlich, and G.E. Strecker. Abstract and concrete categories: The joy of cats. `http://katmat.math.uni-bremen.de/acc`, 2005.

[KS04] V. Kanovei and S. Shelah. A definable nonstandard model of the reals. **Journal of Symbolic Logic**, 69:159–164, 2004.

[Lei13] T. Leinster. Codensity and the ultrafilter monad. Retrieved from arXiv, June 2013.

[Mar02] D. Marker. **Model Theory: An Introduction**. Springer-Verlag Inc., New York, first edition, 2002.

[Mat87] G. Matthiessen. Regular and strongly finitary structures over strongly algebroidal categories. **Canadian Journal of Mathematics**, 30:250–261, 1987.

[ML98] S. Mac Lane. **Categories for the Working Mathematician**. Springer-Verlag, New York, second edition, 1998.

[MR77] M. Makkai and G. E. Reyes. **First Order Categorical Logic: Model-Theoretic Methods in the Theory of Topoi and Related Categories**. Springer-Verlag Inc., Berlin, Heidelberg, New York, first edition, 1977. Includes 2008 errata.

[Okh66] T. Okhuma. Ultrapowers in categories. **Yokohama Mathematics Journal**, 14:17–37, 1966.

[Sho77] J.R. Shoenfield. Axioms of set theory. In J. Barwise, editor, **Handbook of Mathematical Logic**, chapter B1, pages 321–344. North-Holland Publishing Company, Amsterdam, New York, Oxford, first edition, 1977.

[Wil95] J.S. Wilson. On simple pseudofinite groups. **Journal of the London Mathematical Society**, 51(2):471–490, 1995.